



**UNIVERSIDAD CÉSAR VALLEJO**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Sistema de pronóstico geolocalizado de mortalidad**

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:**

**Ingeniero de Sistemas**

**AUTORES:**

Bullon Santana, Aldair (ORCID: 0000-0002-9669-0776)

Perez Alarcon, Jean Brayan Alberto (ORCID: 0000-0003-1002-1932)

**ASESOR:**

Dr. Alfaro Paredes, Emigdio Antonio (ORCID:0000-0002-0309-9195)

**LÍNEA DE INVESTIGACIÓN:**

Sistema de información y comunicaciones

LIMA – PERÚ

2020

### **Dedicatoria**

A Dios por permitirnos nacer en esta época y darnos la salud y la fuerza necesaria para lograr cada uno de nuestros objetivos. A nuestros padres por habernos enseñado a enfrentar siempre las adversidades sin perder nunca la dignidad en el intento y con buenos valores.

### **Agradecimiento**

A nuestro asesor el Dr. Emigdio Antonio Alfaro Paredes por su orientación, motivación y valiosos consejos para la elaboración del proyecto de tesis ya que han sido fundamentales para nuestra formación profesional.

## Índice de contenidos

I. INTRODUCCIÓN .....	1
II. MARCO TEÓRICO .....	12
III. MÉTODO .....	20
<b>3.1 Tipo y diseño de investigación</b> .....	21
<b>3.2 Variables y operacionalización</b> .....	21
<b>3.3 Población, muestra y muestreo</b> .....	22
<b>3.4 Técnicas e instrumentos de recolección de datos</b> .....	23
<b>3.5 Procedimientos</b> .....	24
<b>3.6 Método de análisis de datos</b> .....	25
<b>3.7 Aspectos éticos</b> .....	27
IV. RESULTADOS .....	28
V. DISCUSIÓN .....	52
VI. CONCLUSIONES.....	58
VII. RECOMENDACIONES.....	63
REFERENCIAS.....	66

## Índice de tablas

Tabla 1 ANOVA del modelo de regresión lineal múltiple .....	30
Tabla 2 Resumen del modelo de regresión lineal múltiple .....	30
Tabla 3 Estadísticos descriptivos del modelo de regresión lineal múltiple .....	30
Tabla 4 Coeficientes del modelo de regresión lineal múltiple .....	31
Tabla 5 Resumen de los modelos de regresión lineal múltiple por región .....	38
Tabla 6 Resumen del modelo de regresión logística ordinal .....	39
Tabla 7 Coeficientes del modelo de regresión logística ordinal .....	39
Tabla 8 Probabilidades del modelo de regresión logística ordinal .....	41
Tabla 9 Resumen de los modelos de regresión logística ordinal por región .....	46
Tabla 10 Resumen del modelo ARIMA .....	47
Tabla 11 Coeficientes del modelo ARIMA .....	47
Tabla 12 Resumen de los modelos ARIMA por región .....	49
Tabla 13 Resumen de los resultados de las comprobaciones de las hipótesis .....	51
Tabla 14 Matriz de operacionalización de variables .....	75
Tabla 15 Matriz de consistencia .....	76
Tabla 16 Comparación entre modelos de minería de datos .....	85
Tabla 17 Comparación entre técnicas predictivas de minería de datos .....	86
Tabla 18 Comparación entre metodologías de minería de datos .....	87
Tabla 19 Selección de la metodología de minería de datos .....	88
Tabla 20 Comparación entre herramientas GIS .....	92
Tabla 21 Selección de la herramienta GIS .....	93
Tabla 22 Comprensión del negocio - Inventario de recursos a nivel de hardware .....	98
Tabla 23 Comprensión del negocio - Inventario de recursos a nivel de software .....	99
Tabla 24 Comprensión del negocio - Plan del proyecto .....	101
Tabla 25 Comprensión de los datos - Tabla ESTADO_CIVIL .....	104
Tabla 26 Comprensión de los datos - Tabla TIPO_EDAD .....	104
Tabla 27 Comprensión de los datos - Tabla CALIDAD_MEDICA .....	105
Tabla 28 Comprensión de los datos - Tabla NIVEL_INSTRUCCION .....	105
Tabla 29 Comprensión de los datos - Tabla TIPO_ZONA .....	105
Tabla 30 Comprensión de los datos - Tabla LOCAL_DEFUNCION .....	106
Tabla 31 Comprensión de los datos - Tabla FUNDAMENTO_CAUSA .....	106
Tabla 32 Comprensión de los datos - Tabla ESTADO_NUTRICIONAL .....	106
Tabla 33 Comprensión de los datos - Tabla REGIÓN .....	107
Tabla 34 Comprensión de los datos - Tabla PROVINCIA .....	107
Tabla 35 Comprensión de los datos - Tabla COMUNA .....	107
Tabla 36 Comprensión de los datos - Tabla SERVICIO_SALUD .....	108
Tabla 37 Comprensión de los datos - Tabla ACTIVIDAD .....	108
Tabla 38 Comprensión de los datos - Tabla CATEGORIA_OCUPACIONAL .....	108
Tabla 39 Comprensión de los datos - Tabla OCUPACION .....	109
Tabla 40 Comprensión de los datos - Tabla PERSONA .....	109
Tabla 41 Comprensión de los datos - Tabla PERSONA_FAMILIARES .....	110
Tabla 42 Comprensión de los datos - Tabla CERTIFICADO_DEFUNCION .....	111
Tabla 43 Comprensión de los datos - Tabla CIE10_CAPITULO .....	112
Tabla 44 Comprensión de los datos - Tabla CIE10_GRUPO .....	112
Tabla 45 Comprensión de los datos - Tabla CIE10_CATEGORIA .....	112
Tabla 46 Comprensión de los datos - Tabla CIE10_ENFERMEDAD .....	113
Tabla 47 Cantidad de muertes por sexo y región .....	114
Tabla 48 Cantidad de muertes por estado civil y región .....	115

Tabla 49 Cantidad de muertes por tipo de zona y región .....	117
Tabla 50 Cantidad de muertes por actividad y región .....	118
Tabla 51 Cantidad de muertes por capítulo de CIE-10 y región .....	121
Tabla 52 Cantidad de muertes por año y región .....	124
Tabla 53 Cantidad de muertes por autolesiones y región del año 1997 .....	126
Tabla 54 Cantidad de muertes mayores a 200 de adolescentes en Santiago .....	127
Tabla 55 Top 10 cantidad de muertes por región desde el año 2000 al 2017 .....	128
Tabla 56 Principales 5 causas de muertes en Santiago entre el año 2000 y 2017 ....	129
Tabla 57 Cantidad de muertes desde 1997 a 2017 .....	130
Tabla 58 Cantidad de muertes por estado nutricional de niños menores de 1 año ...	131
Tabla 59 Edad promedio de los fallecidos por categoría ocupacional del 2000 al 2017 .....	132
Tabla 60 Edad promedio de los fallecidos por regiones .....	134
Tabla 61 Las 5 regiones/provincias con la mortalidad más baja en enero de 2016...	135
Tabla 62 Preparación de los datos - Campos relevantes para el análisis.....	136
Tabla 63 Preparación de los datos - Problemas de datos .....	138
Tabla 64 Preparación de los datos – Dimensión A00_B99.....	148
Tabla 65 Preparación de los datos - Dimensión C00_D48 .....	148
Tabla 66 Preparación de los datos - Dimensión D50_D89.....	148
Tabla 67 Preparación de los datos - Dimensión E00_E90 .....	149
Tabla 68 Preparación de los datos - Dimensión F00_F99.....	149
Tabla 69 Preparación de los datos - Dimensión G00_G99.....	149
Tabla 70 Preparación de los datos - Dimensión H00_H59.....	150
Tabla 71 Preparación de los datos - Dimensión H60_H95.....	150
Tabla 72 Preparación de los datos - Dimensión I00_I99 .....	150
Tabla 73 Preparación de los datos - Dimensión J00_J99.....	151
Tabla 74 Preparación de los datos - Dimensión K00_K93 .....	151
Tabla 75 Preparación de los datos - Dimensión L00_L99 .....	151
Tabla 76 Preparación de los datos - Dimensión M00_M99 .....	152
Tabla 77 Preparación de los datos - Dimensión N00_N99.....	152
Tabla 78 Preparación de los datos - Dimensión O00_O99.....	152
Tabla 79 Preparación de los datos - Dimensión P00_P96 .....	153
Tabla 80 Preparación de los datos - Dimensión Q00_Q99.....	153
Tabla 81 Preparación de los datos - Dimensión R00_R99.....	153
Tabla 82 Preparación de los datos - Dimensión S00_T98.....	154
Tabla 83 Preparación de los datos - Dimensión SEX.....	154
Tabla 84 Preparación de los datos - Dimensión DEATH_LOCATION .....	154
Tabla 85 Preparación de los datos - Dimensión DIM_AGE_0_10 .....	155
Tabla 86 Preparación de los datos - Dimensión DIM_AGE_11_19 .....	155
Tabla 87 Preparación de los datos - Dimensión DIM_AGE_20_64 .....	155
Tabla 88 Preparación de los datos - Dimensión DIM_AGE_65.....	156
Tabla 89 Preparación de los datos - Dimensión TERRITORIAL_LOCATION.....	156
Tabla 90 Preparación de los datos - Dimensión TIME_PERIOD.....	157
Tabla 91 Preparación de los datos - Tabla de hechos .....	157
Tabla 92 Evaluación - Resumen del modelo de RLM de Valparaíso.....	169
Tabla 93 Evaluación - ANOVA del modelo de RLM de Valparaíso.....	170
Tabla 94 Evaluación – Estadísticos descriptivos del modelo de Valparaíso.....	170
Tabla 95 Evaluación - Coeficientes del modelo de RLM de Valparaíso.....	171
Tabla 96 Evaluación - Resumen del modelo de RLM de Santiago.....	178
Tabla 97 Evaluación - ANOVA del modelo de RLM de Santiago.....	179
Tabla 98 Evaluación - Estadísticos descriptivos del modelo de Santiago.....	179

Tabla 99 Evaluación - Coeficientes del modelo de RLM de Santiago .....	180
Tabla 100 Evaluación - Resumen del modelo de RLMde Biobío.....	187
Tabla 101 Evaluación - ANOVA del modelo de RLM de Biobío .....	188
Tabla 102 Evaluación - Estadísticos descriptivos del modelo de Biobío.....	188
Tabla 103 Evaluación - Coeficientes del modelo de RLM de Biobío .....	189
Tabla 104 Evaluación - Resumen del modelo de RLM de Maule .....	196
Tabla 105 Evaluación - ANOVA del modelo de RLM de Maule .....	197
Tabla 106 Evaluación - Estadísticos descriptivos del modelo de Maule .....	197
Tabla 107 Evaluación - Coeficientes del modelo de RLM de Maule .....	198
Tabla 108 Evaluación - Resumen del modelo de RLM de O'Higgins .....	205
Tabla 109 Evaluación - ANOVA del modelo de RLM de O'Higgins .....	206
Tabla 110 Evaluación - Estadísticos descriptivos del modelo de O'Higgins .....	206
Tabla 111 Evaluación - Coeficientes del modelo de RLM de O'Higgins .....	207
Tabla 112 Evaluación - Resumen del modelo de RLM de Araucanía.....	214
Tabla 113 Evaluación - ANOVA del modelo de RLM de Araucanía .....	215
Tabla 114 Evaluación - Estadísticos descriptivos del modelo de Araucanía.....	215
Tabla 115 Evaluación - Coeficientes del modelo de RLM de Araucanía .....	216
Tabla 116 Evaluación - Resumen del modelo de RLM de Coquimbo.....	223
Tabla 117 Evaluación - ANOVA del modelo de RLM de Coquimbo .....	224
Tabla 118 Evaluación - Estadísticos descriptivos del modelo de Coquimbo.....	224
Tabla 119 Evaluación - Coeficientes del modelo de RLM de Coquimbo .....	225
Tabla 120 Evaluación - Resumen del modelo de RLM de Ñuble.....	232
Tabla 121 Evaluación - ANOVA del modelo de RLM de Ñuble .....	233
Tabla 122 Evaluación - Estadísticos descriptivos del modelo de Ñuble .....	233
Tabla 123 Evaluación - Coeficientes del modelo de RLM de Ñuble .....	234
Tabla 124 Evaluación - Resumen del modelo de RLM de Antofagasta .....	241
Tabla 125 Evaluación - ANOVA del modelo de RLM de Antofagasta.....	242
Tabla 126 Evaluación - Estadísticos descriptivos del modelo de Antofagasta.....	242
Tabla 127 Evaluación - Coeficientes del modelo de RLM de Antofagasta.....	243
Tabla 128 Evaluación - Resumen del modelo de RLM de Los Lagos .....	250
Tabla 129 Evaluación - ANOVA del modelo de RLM de Los Lagos .....	251
Tabla 130 Evaluación - Estadísticos descriptivos del modelo de Los Lagos .....	251
Tabla 131 Evaluación - Coeficientes del modelo de RLM de Los Lagos .....	252
Tabla 132 Evaluación - Resumen del modelo de RLM de Los Ríos.....	259
Tabla 133 Evaluación - ANOVA del modelo de RLM de Los Ríos.....	260
Tabla 134 Evaluación - Estadísticos descriptivos del modelo de Los Ríos.....	260
Tabla 135 Evaluación - Coeficientes del modelo de RLM de Los Ríos.....	261
Tabla 136 Evaluación - Resumen del modelo de RLM de Atacama.....	268
Tabla 137 Evaluación - ANOVA del modelo de RLM de Atacama.....	269
Tabla 138 Evaluación - Estadísticos descriptivos del modelo de Atacama.....	269
Tabla 139 Evaluación - Coeficientes del modelo de RLM de Atacama.....	270
Tabla 140 Evaluación - Resumen del modelo de RLM de Tarapacá.....	277
Tabla 141 Evaluación - ANOVA del modelo de RLM de Tarapacá.....	278
Tabla 142 Evaluación - Estadísticos descriptivos del modelo de Tarapacá.....	278
Tabla 143 Evaluación - Coeficientes del modelo de RLM de Tarapacá.....	279
Tabla 144 Evaluación - Resumen del modelo de RLM de Arica y Parinacota .....	286
Tabla 145 Evaluación - ANOVA del modelo de RLM de Arica y Parinacota .....	287
Tabla 146 Evaluación - Estadísticos descriptivos del modelo de Arica y Parinacota .....	287
Tabla 147 Evaluación - Coeficientes del modelo de RLM de Arica y Parinacota .....	288
Tabla 148 Evaluación - Resumen del modelo de RLM de Magallanes.....	294
Tabla 149 Evaluación - ANOVA del modelo de RLM de Magallanes .....	295

Tabla 150 Evaluación - Estadísticos descriptivos del modelo de Magallanes.....	295
Tabla 151 Evaluación - Coeficientes del modelo de RLM de Magallanes .....	296
Tabla 152 Evaluación - Resumen del modelo de RLM de Aysen .....	303
Tabla 153 Evaluación - ANOVA del modelo de RLM de Aysen.....	304
Tabla 154 Evaluación - Estadísticos descriptivos del modelo de Aysen.....	304
Tabla 155 Evaluación - Coeficientes del modelo de RLM de Aysen.....	305
Tabla 156 Evaluación - Resumen de los modelos de RLM por cada región.....	312
Tabla 157 Evaluación - Coeficientes del modelo de RLO de Santiago.....	313
Tabla 158 Evaluación - Probabilidades del modelo de RLO de Santiago.....	314
Tabla 159 Evaluación - Coeficientes del modelo de RLO de Valparaíso.....	315
Tabla 160 Evaluación - Probabilidades del modelo de RLO de Valparaíso.....	316
Tabla 161 Evaluación - Coeficientes del modelo de RLO de Biobío.....	317
Tabla 162 Evaluación - Probabilidades del modelo de RLO de Biobío.....	318
Tabla 163 Evaluación - Coeficientes del modelo de RLO de Araucanía.....	319
Tabla 164 Evaluación - Probabilidades del modelo de RLO de Araucanía.....	320
Tabla 165 Evaluación - Coeficientes del modelo de RLO de Maule .....	321
Tabla 166 Evaluación - Probabilidades del modelo de RLO de Maule .....	322
Tabla 167 Evaluación - Coeficientes del modelo de RLO de O'Higgins .....	323
Tabla 168 Evaluación - Probabilidades del modelo de RLO de O'Higgins .....	324
Tabla 169 Evaluación - Coeficientes del modelo de RLO de Los Lagos .....	325
Tabla 170 Evaluación - Probabilidades del modelo de RLO de Los Lagos .....	326
Tabla 171 Evaluación - Coeficientes del modelo de RLO de Coquimbo.....	327
Tabla 172 Evaluación - Probabilidades del modelo de RLO de Coquimbo.....	328
Tabla 173 Evaluación - Coeficientes del modelo de RLO de Ñuble .....	329
Tabla 174 Evaluación - Probabilidades del modelo de RLO de Ñuble .....	330
Tabla 175 Evaluación - Coeficientes del modelo de RLO de Antofagasta.....	331
Tabla 176 Evaluación - Probabilidades del modelo de RLO de Antofagasta.....	332
Tabla 177 Evaluación - Coeficientes del modelo de RLO de Los Ríos.....	333
Tabla 178 Evaluación - Probabilidades del modelo de RLO de Los Ríos.....	334
Tabla 179 Evaluación - Coeficientes del modelo de RLO de Atacama.....	335
Tabla 180 Evaluación - Probabilidades del modelo de RLO de Atacama.....	336
Tabla 181 Evaluación - Coeficientes del modelo de RLO de Tarapacá.....	337
Tabla 182 Evaluación - Probabilidades del modelo de RLO de Tarapacá.....	338
Tabla 183 Evaluación - Coeficientes del modelo de RLO de Arica y Parinacota .....	339
Tabla 184 Evaluación - Probabilidades del modelo de RLO de Arica y Parinacota ...	340
Tabla 185 Evaluación - Coeficientes del modelo de RLO de Magallanes.....	341
Tabla 186 Evaluación - Probabilidades del modelo de RLO de Magallanes.....	342
Tabla 187 Evaluación - Coeficientes del modelo de RLO de Aysen.....	343
Tabla 188 Evaluación - Probabilidades del modelo de RLO de Aysen .....	344
Tabla 189 Evaluación - Resumen del modelo ARIMA de Santiago .....	345
Tabla 190 Evaluación - Predicciones estimadas para la región de Santiago.....	346
Tabla 191 Evaluación - Resumen del modelo ARIMA de Valparaíso .....	347
Tabla 192 Evaluación - Predicciones estimadas para la región de Valparaíso.....	348
Tabla 193 Evaluación - Resumen del modelo ARIMA de Biobío.....	349
Tabla 194 Evaluación - Predicciones estimadas para la región de Biobío.....	350
Tabla 195 Evaluación - Resumen del modelo ARIMA de Maule .....	351
Tabla 196 Evaluación - Predicciones estimadas para la región de Maule .....	352
Tabla 197 Evaluación - Resumen del modelo ARIMA de Araucanía.....	353
Tabla 198 Evaluación - Predicciones estimadas para la región de Araucanía.....	354
Tabla 199 Evaluación - Resumen del modelo ARIMA de O'Higgins.....	355
Tabla 200 Evaluación - Predicciones estimadas para la región de O'Higgins .....	356



Tabla 201 Evaluación - Resumen del modelo ARIMA de Los Lagos.....	357
Tabla 202 Evaluación - Predicciones estimadas para la región de Los Lagos .....	358
Tabla 203 Evaluación - Resumen del modelo ARIMA de Coquimbo.....	359
Tabla 204 Evaluación - Predicciones estimadas para la región de Coquimbo.....	360
Tabla 205 Evaluación - Resumen del modelo ARIMA de Ñuble.....	361
Tabla 206 Evaluación - Predicciones estimadas para la región de Ñuble .....	362
Tabla 207 Evaluación - Resumen del modelo ARIMA de Antofagasta .....	363
Tabla 208 Evaluación - Predicciones estimadas para la región de Antofagasta.....	364
Tabla 209 Evaluación - Resumen del modelo ARIMA de Los Ríos.....	365
Tabla 210 Evaluación - Predicciones estimadas para la región de Los Ríos.....	366
Tabla 211 Evaluación - Resumen del modelo ARIMA de Atacama .....	367
Tabla 212 Evaluación - Predicciones estimadas para la región de Atacama.....	368
Tabla 213 Evaluación - Resumen del modelo ARIMA de Tarapacá.....	369
Tabla 214 Evaluación - Predicciones estimadas para la región de Tarapacá.....	370
Tabla 215 Evaluación - Resumen del modelo ARIMA de Arica y Parinacota .....	371
Tabla 216 Evaluación - Predicciones estimadas para la región Arica y Parinacota ...	372
Tabla 217 Evaluación - Resumen del modelo ARIMA de Magallanes.....	373
Tabla 218 Evaluación - Predicciones estimadas para la región de Magallanes.....	374
Tabla 219 Evaluación - Resumen del modelo ARIMA de Aysén .....	375
Tabla 220 Evaluación - Predicciones estimadas para la región de Aysén.....	376
Tabla 221 Evaluación - Resumen de los modelos ARIMA por cada región .....	377

## Índice de figuras

<i>Figura 1.</i> Ecuación de RLM, adaptado de "A Multiple Linear Regression Model for structure of N-Linked oligosaccharides", por Cheng et al., 2015, the institution of Engineering & Technology.....	26
<i>Figura 2.</i> Ecuación de RLO, adaptado de "A Multiple Linear Regression Model for structure of N-Linked oligosaccharides",por Cheng et al., 2015, The institution of Engineering & Technology.....	26
<i>Figura 3.</i> Ecuación del modelo ARIMA, adaptado de "Temporal analysis of road accidents by ARIMA model", por Gherida et al., 2018, International Journal of Innovation and Applied Studie.....	27
<i>Figura 4.</i> Comportamiento de la mortalidad de Santiago .....	46
<i>Figura 5.</i> Residuos estandarizados del modelo ARIMA de Santiago.....	47
<i>Figura 6.</i> ACF de los residuos del modelo ARIMA de Santiago .....	48
<i>Figura 7.</i> Prueba del ruido blanco del modelo ARIMA de Santiago .....	48
<i>Figura 8.</i> Valores precedidos del modelo ARIMA vs. valores esperados.....	48
<i>Figura 9.</i> Pronóstico del modelo ARIMA .....	49
<i>Figura 10.</i> Comprensión de los datos - Recolección de datos iniciales .....	102
<i>Figura 11.</i> Comprensión de los datos - Diagrama de la base de datos relacional .....	103
<i>Figura 12.</i> Consulta en la base de datos de muertes por sexo y región .....	113
<i>Figura 13.</i> Cantidad de muertes por sexo y región.....	114
<i>Figura 14.</i> Consulta en la base de datos de muertes por estado civil y región .....	115
<i>Figura 15.</i> Cantidad de muertes por estado civil y región.....	116
<i>Figura 16.</i> Consulta en la base de datos de cantidad de muertes por tipo de zona y región.....	116
<i>Figura 17.</i> Cantidad de muertes por tipo de zona y región .....	117
<i>Figura 18.</i> Consulta en la base de datos de cantidad de muertes por actividad y región .....	118
<i>Figura 19.</i> Cantidad de muertes por actividad y región .....	119
<i>Figura 20.</i> Consulta en la base de datos de cantidad de muertes por capítulo CIE-10 y región.....	120
<i>Figura 21.</i> Cantidad de muertes por capítulo CIE-10 y región.....	122
<i>Figura 22.</i> Consulta en la base de datos de cantidad de muertes por año y región... ..	123
<i>Figura 23.</i> Cantidad de muertes por año y región .....	125
<i>Figura 24.</i> Consulta en la base de datos de cantidad de muertes por autolesiones y región del año 1997 .....	126
<i>Figura 25.</i> Cantidad de muertes por autolesiones y región del año 1997 .....	126
<i>Figura 26.</i> Consulta en la base de datos de cantidad de muertes mayores a 200 de adolescentes en Santiago.....	127
<i>Figura 27.</i> Cantidad de muertes mayores a 200 de adolescentes en Santiago .....	127
<i>Figura 28.</i> Consulta en la base de datos del top 10 de la cantidad de muertes por región desde el año 200 al 2017 .....	128
<i>Figura 29.</i> Cantidad de muertes por región desde el año 2000 al 2017 .....	128
<i>Figura 30.</i> Consulta en la base de datos de las 5 causas de muerte en Santiago entre el año 2000 y 2017 .....	129
<i>Figura 31.</i> Las 5 causas de muerte en Santiago entre el año 2000 y 2017 .....	129
<i>Figura 32.</i> Consulta en la base de datos de cantidad de muertes desde el año 1997 al 2017 .....	130
<i>Figura 33.</i> Cantidad de muertes desde el año 1997 al 2017 .....	130

<i>Figura 34.</i> Consulta en la base de datos de cantidad de muertes por estado nutricional de niños menores a 1 año.....	131
<i>Figura 35.</i> Cantidad de muertes por estado nutricional de niños menores a 1 año ...	131
<i>Figura 36.</i> Consulta en la base de datos de las edades promedio de los fallecidos por categoría ocupacional del año 2000 al 2017 .....	132
<i>Figura 37.</i> Edades promedio de fallecidos por categoría ocupacional.....	132
<i>Figura 38.</i> Edades promedio de fallecidos por categoría ocupacional.....	133
<i>Figura 39.</i> Consulta en la base de datos de edad promedio de los fallecidos por regiones.....	133
<i>Figura 40.</i> Edades promedio de los fallecidos por regiones .....	134
<i>Figura 41.</i> Consulta en la base de datos de las 5 regiones/provincias con la mortalidad más baja en enero del 2016.....	135
<i>Figura 42.</i> Las 5 regiones/provincias con la mortalidad más baja en enero 2016.....	135
<i>Figura 43.</i> Preparación de los datos - Consulta SQL para obtener los datos ambiguos .....	138
<i>Figura 44.</i> Preparación de los datos - Resultados de los datos ambiguos .....	139
<i>Figura 45.</i> Preparación de los datos - Consulta SQL para obtener los códigos inconsistentes .....	139
<i>Figura 46.</i> Preparación de los datos - Códigos inconsistentes .....	139
<i>Figura 47.</i> Preparación de los datos - Consulta SQL para actualizar los códigos.....	140
<i>Figura 48.</i> Preparación de los datos - Resultados de los códigos actualizados.....	140
<i>Figura 49.</i> Preparación de los datos - Consulta SQL para visualizar los datos perdidos .....	140
<i>Figura 50.</i> Preparación de los datos - Datos perdidos.....	141
<i>Figura 51.</i> Preparación de los datos - Validar datos perdidos .....	141
<i>Figura 52.</i> Preparación de los datos - Resultados de la validación de datos perdidos .....	141
<i>Figura 53.</i> Preparación de los datos - Consulta SQL para actualizar los datos perdidos .....	142
<i>Figura 54.</i> Preparación de los datos - Script para la creación del modelo estrella ....	146
<i>Figura 55.</i> Preparación de los datos - Diagrama del modelo estrella .....	147
<i>Figura 56.</i> Preparación de los datos – Script para cargar los datos al modelo dimensional .....	161
<i>Figura 57.</i> Modelado - Interfaz de Azure ML .....	163
<i>Figura 58.</i> Modelado - Importar datos a Azure ML .....	163
<i>Figura 59.</i> Modelado - Conectar la base de datos con Azure ML .....	164
<i>Figura 60.</i> Modelado - Script para crear el modelo de regresión lineal múltiple.....	165
<i>Figura 61.</i> Modelado - Script para crear el modelo de regresión logística ordinal ....	167
<i>Figura 62.</i> Modelado - Script para crear el modelo ARIMA .....	168
<i>Figura 63.</i> Modelado - Modelo en Azure ML .....	168
<i>Figura 64.</i> Evaluación - Comportamiento de la mortalidad de Santiago .....	345
<i>Figura 65.</i> Evaluación - Ajuste del modelo ARIMA de Santiago .....	345
<i>Figura 66.</i> Evaluación - Pronóstico de la mortalidad de Santiago.....	346
<i>Figura 67.</i> Evaluación - Comportamiento de la mortalidad de Valparaíso .....	347
<i>Figura 68.</i> Evaluación - Ajuste del modelo ARIMA de Valparaíso .....	347
<i>Figura 69.</i> Evaluación - Pronóstico de la mortalidad de Valparaíso.....	348
<i>Figura 70.</i> Evaluación - Comportamiento de la mortalidad de Biobío .....	349
<i>Figura 71.</i> Evaluación - Ajuste del modelo ARIMA de Biobío .....	349
<i>Figura 72.</i> Evaluación - Pronóstico de la mortalidad de Biobío .....	350
<i>Figura 73.</i> Evaluación - Comportamiento de la mortalidad de Maule .....	351
<i>Figura 74.</i> Evaluación - Ajuste del modelo ARIMA de Maule .....	351

<i>Figura 75.</i> Evaluación - Pronóstico de la mortalidad de Maule .....	352
<i>Figura 76.</i> Evaluación - Comportamiento de la mortalidad de Araucanía .....	353
<i>Figura 77.</i> Evaluación - Ajuste del modelo ARIMA de Araucanía .....	353
<i>Figura 78.</i> Evaluación - Pronóstico de la mortalidad de Araucanía .....	354
<i>Figura 79.</i> Evaluación - Comportamiento de la mortalidad de O'Higgins .....	355
<i>Figura 80.</i> Evaluación - Ajuste del modelo ARIMA de O'Higgins .....	355
<i>Figura 81.</i> Evaluación - Pronóstico de la mortalidad de O'Higgins .....	356
<i>Figura 82.</i> Evaluación - Comportamiento de la mortalidad de Los Lagos .....	357
<i>Figura 83.</i> Evaluación - Ajuste del modelo ARIMA de Los Lagos .....	35
<i>Figura 84.</i> Evaluación - Pronóstico de la mortalidad de Los Lagos .....	358
<i>Figura 85.</i> Evaluación - Comportamiento de la mortalidad de Coquimbo .....	359
<i>Figura 86.</i> Evaluación - Ajuste del modelo ARIMA de Coquimbo .....	359
<i>Figura 87.</i> Evaluación - Pronóstico de la mortalidad de Coquimbo .....	360
<i>Figura 88.</i> Evaluación - Comportamiento de la mortalidad de Ñuble .....	361
<i>Figura 89.</i> Evaluación - Ajuste del modelo ARIMA de Ñuble .....	361
<i>Figura 90.</i> Evaluación - Pronóstico de la mortalidad de Ñuble .....	362
<i>Figura 91.</i> Evaluación - Comportamiento de la mortalidad de Antofagasta .....	363
<i>Figura 92.</i> Evaluación - Ajuste del modelo ARIMA de Antofagasta .....	363
<i>Figura 93.</i> Evaluación - Pronóstico de la mortalidad de Antofagasta .....	364
<i>Figura 94.</i> Evaluación - Comportamiento de la mortalidad de Los Ríos .....	365
<i>Figura 95.</i> Evaluación - Ajuste del modelo ARIMA de Los Ríos .....	365
<i>Figura 96.</i> Evaluación - Pronóstico de la mortalidad de Los Ríos .....	366
<i>Figura 97.</i> Evaluación - Comportamiento de la mortalidad de Atacama .....	367
<i>Figura 98.</i> Evaluación - Ajuste del modelo ARIMA de Atacama .....	367
<i>Figura 99.</i> Evaluación - Pronóstico de la mortalidad de Atacama .....	368
<i>Figura 100.</i> Evaluación - Comportamiento de la mortalidad de Tarapacá .....	369
<i>Figura 101.</i> Evaluación - Ajuste del modelo ARIMA de Tarapacá .....	369
<i>Figura 102.</i> Evaluación - Pronóstico de la mortalidad de Tarapacá .....	370
<i>Figura 103.</i> Evaluación - Comportamiento de la mortalidad de Arica y Parinacota .....	371
<i>Figura 104.</i> Evaluación - Ajuste del modelo ARIMA de Arica y Parinacota .....	371
<i>Figura 105.</i> Evaluación - Pronóstico de la mortalidad de Arica y Parinacota .....	372
<i>Figura 106.</i> Evaluación - Comportamiento de la mortalidad de Magallanes .....	373
<i>Figura 107.</i> Evaluación - Ajuste del modelo ARIMA de Magallanes .....	373
<i>Figura 108.</i> Evaluación - Pronóstico de la mortalidad de Magallanes .....	374
<i>Figura 109.</i> Evaluación - Comportamiento de la mortalidad de Aysén .....	375
<i>Figura 110.</i> Evaluación - Ajuste del modelo ARIMA de Aysén .....	375
<i>Figura 111.</i> Evaluación - Pronóstico de la mortalidad de Aysén .....	376
<i>Figura 112.</i> Implementación - Interfaz del Login .....	379
<i>Figura 113.</i> Implementación - Interfaz del Dashboard 1 .....	380
<i>Figura 114.</i> Implementación - Interfaz del Dashboard 2 .....	381

## Índice de anexos

Anexo 1: Matriz de operacionalización de variables .....	90
Anexo 2: Matriz de consistencia.....	91
Anexo 3: Modelos, técnicas y metodologías de minería de datos .....	92
Anexo 4: Sistema de información geográfica .....	104
Anexo 5: Clasificación internacional de enfermedades .....	109
Anexo 6: Desarrollo de la metodología CRISP-DM.....	111

## Índice de abreviaturas

Abreviatura	Significado	Página
<b>ANOVA</b>	Analysis of the Variance (Análisis de varianza)	46
<b>AR</b>	Autoregressive (Autoregresivo)	185
<b>ARIMA</b>	Autoregressive Integrated Moving Average (Media móvil integrada autorregresiva)	25
<b>AZURE ML</b>	Azure Machine Learning	115
<b>CIE</b>	Clasificación internacional de enfermedades	36
<b>CRISP-DM</b>	Cross Industry Standard Process for Data Mining (Proceso estándar entre industrias para la minería de datos)	33
<b>DEIS</b>	Departamento de Estadísticas e Información de la Salud	39
<b>DM</b>	Data Mining (Minería de datos)	21
<b>GIS</b>	Geographic Information System (Sistema de información geográfica)	20
<b>IQ</b>	Intelligence Quotient (Coeficiente intelectual)	145
<b>KDD</b>	Knowledge Discovery in Databases (Descubrimiento de conocimiento en bases de datos)	21
<b>LR</b>	Linear Regression (Regresión lineal)	42
<b>MA</b>	Movil Average (Media móvil)	150
<b>MAE</b>	Mean Absolute Error (Error absoluto medio)	63
<b>MAPE</b>	Mean Absolute Percentage Error (Error de porcentaje absoluto medio)	26
<b>OMS</b>	Organización Mundial de la Salud	89
<b>RLM</b>	Regresión lineal múltiple	42
<b>RLO</b>	Regresión logística ordinal	42
<b>RMSE</b>	Root Mean Square Error (Error cuadrático medio)	63
<b>SEMMA</b>	Sampling, Exploring, Modifying, Modeling, and Assesing (Muestreo, exploración, modificación, modelado y evaluación)	33
<b>SARIMA</b>	Seasonal Autoregressive Integrated Moving Average (Media móvil integrada autorregresiva estacional)	26
<b>VIF</b>	Variance Inflation Factor (Factor de inflación de la varianza)	47

## Resumen

El problema de la investigación fue: ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo? El objetivo de la investigación fue determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo. El tipo de investigación fue aplicada y el tipo de diseño fue no experimental, longitudinal y de tendencias debido a que se analizaron los datos de mortalidad de Chile en el período 1997-2017. Los datos son de acceso público y fueron obtenidos del Departamento de Estadística e Información de Salud. Para el desarrollo de la investigación se usó un proceso de minería de datos aplicando la metodología CRISP-DM. Las técnicas usadas para el análisis de datos fueron: regresión lineal múltiple, regresión logística ordinal y series de tiempo.

Los resultados obtenidos en la presente investigación demostraron que con la técnica de regresión lineal múltiple se obtuvo una precisión mayor al 80% en al menos dos regiones de Chile. La técnica de regresión logística ordinal tuvo una precisión mayor al 66% en al menos dos regiones de Chile. La técnica de series de tiempo usando el modelo ARIMA obtuvo un MAPE menor al 5% en al menos tres regiones de Chile. Con los resultados obtenidos en la investigación se llegó a concluir que existen causas de muertes que tienen un impacto superior al 70% en al menos una región de Chile.

Finalmente, se recomendó aplicar otros modelos de pronóstico considerando una mayor especificidad en la CIE-10, debido a que la presente investigación se centró en el análisis por capítulos, pero cada capítulo tiene grandes grupos de enfermedades. Esto ayudará a tener un análisis detallado o un modelo que identifique el patrón de mortalidad para cada grupo, categoría o enfermedad. También, se sugirió llevar a cabo investigaciones en un área geográfica mayor a la utilizada en la investigación. Por otro lado, se recomendó el uso de Machine Learning y Deep Learning, ya que podrían tener mejores resultados.

**Palabras clave:** Mortalidad, Series de tiempo, Regresión lineal múltiple, Regresión logística ordinal, Pronóstico

## Abstract

The problem of the investigation was: What was the improvement in the precision of the mortality forecast with the multiple linear regression technique, ordinal logistic regression and time series? The objective of the study was to determine the precision of the mortality prognosis with the multiple linear regression technique, ordinal logistic regression and time series. The type of the study was applied and the type of design was non-experimental, longitudinal and trends because the mortality data from Chile were analyzed in the period 1997-2017. The data are publicly accessible and were obtained from the Department of Health Statistics and Information. For the development of the study, a data mining process was used applying the CRISP-DM methodology. The techniques used for data analysis were: multiple linear regression, ordinal logistic regression, and time series.

The results obtained in the present investigation showed that with the multiple linear regression technique a precision greater than 80% was obtained in at least two regions of Chile. The ordinal logistic regression technique had a precision greater than 66% in at least two regions of Chile. The time series technique using the ARIMA model obtained a MAPE of less than 5% in at least three regions of Chile. With the results obtained in the investigation, it was concluded that there are causes of death that have an impact greater than 70% in at least one region of Chile.

Finally, it was recommended to apply other forecasting models considering a greater specificity in the CIE-10, because the present investigation focused on the analysis by chapters, but each chapter has large groups of diseases. This will help to have a detailed analysis or model that identifies the mortality pattern for each group, category or disease. It was also suggested to carry out research in a geographic area larger than that used in the research. On the other hand, it was advised that the use of Machine Learning and Deep Learning could have better results.

**Keywords:** Mortality, Time Series, Multiple Linear Regression, Ordinal Logistic Regression, Forecast



# **I. INTRODUCCIÓN**

Es importante tener una proyección correcta acerca de las causas de muerte más comunes en el país, debido a que estos datos ayudan a los gobiernos a planificar actividades en el campo de la salud pública con el fin de evitar más decesos. Por lo cual es fundamental tener un modelo de pronóstico en base a cada región de un País para tener una mejor precisión. Se planteó como problema general: ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo?. El objetivo general fue determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo. Esta investigación se justificó de manera teórica y económica porque gracias a ello se podrá crear estrategias para reducir la mortalidad y tener una inversión correcta para evitar enfermedades. Como hipótesis general se planteó que el uso de técnicas de minería de datos aumenta la precisión y reduce el error porcentual absoluto medio del pronóstico de mortalidad.

La mortalidad es actualmente uno de los principales problemas a prevenir y analizar para apoyar el desarrollo y planificación de la salud pública. Al respecto, Arrieta, Flórez y Alvis (2018) explicaron: “Hoy la muerte es un hecho inevitable, sin embargo, la esperanza de vida ha cambiado significativamente con el tiempo y todavía difiere ampliamente entre los grupos humanos de hoy” (p. 1). y que “El avance de la tecnología sanitaria ha permitido controlar, retrasar o evitar ciertas muertes” (p. 1). Por otro lado, Kumar y Garg (2013) indicaron:

La información sobre las causas de muerte es muy importante para la formulación de políticas, planificación, monitoreo y evaluación de programas de salud, además de ser necesaria para la investigación de campo, las comparaciones y el conocimiento de la epidemia. En los países en desarrollo, donde la mayoría de las muertes no son atendidas por un médico ni certificadas médicamente, esta información crucial a menudo es incompleta y de mala calidad. (p. 3)

Por lo mencionado anteriormente, esto lleva a preguntarse: ¿Por qué es de suma importancia conocer las causas de muertes? Los especialistas de la Organización Mundial de la Salud (s. f.) señalaron:

Las estadísticas sobre las causas de muerte ayudan a las autoridades sanitarias a orientar futuras actividades en el campo de la salud pública. Por ejemplo, cuando la mortalidad por enfermedades del corazón y diabetes se dispara en unos pocos años, es altamente deseable lanzar un programa poderoso para promover estilos de vida que ayuden a prevenir estas enfermedades. Del mismo modo, si hay una alta tasa de mortalidad infantil por neumonía en un país y el presupuesto dedicado a ofrecer tratamientos efectivos es pequeño, se puede aumentar el gasto en esta área. (párr. 10)

En base a lo mencionado por los especialistas de la Organización Mundial de la Salud (OMS), se concluye que es muy importante predecir las causas de muerte para reducir la tasa de mortalidad. Al respecto, Giraldo, Atehortúa, García, Díaz, Romero y Rodríguez (2017) sustentaron:

El monitoreo y la predicción de la mortalidad son indicadores ampliamente utilizados por los sistemas de salud para tomar decisiones informadas, formular políticas y directrices en el sector de la salud, priorizar el gasto y determinar la carga de la enfermedad. Debido a la complejidad asociada con las relaciones entre las diferentes variables que determinan la mortalidad en una población, la mayoría de los análisis se llevan a cabo utilizando algún modelo que establece alguna relación entre los datos. (p. 248)

Por eso es cada vez más esencial contar con nuevas técnicas y herramientas que faciliten la forma de predecir la cantidad de muertes en un lugar determinado. Al respecto, Hernández, Díaz, Espinoza y Vilcarromero (2016) explicaron:

El análisis geoespacial es extremadamente importante para abordar los problemas de salud pública y la provisión rápida de servicios de salud y su implementación en programas o estrategias dirigidos a la población más afectada, para tener recursos humanos y logísticos disponibles más cerca para disminuir o eliminar el impacto de las enfermedades analizadas. (p. 186)

Por lo tanto, una de las herramientas emergentes que puede ayudar a resolver este problema es el Sistema de Información Geográfica, conocido como SIG o por sus siglas en inglés GIS (*Geographic Information System*). Sobre los sistemas de información geográfica, Wang, Du y Xie (2009) indicaron:

GIS es una tecnología para la recopilación, almacenamiento, consulta y análisis de información espacial sintética, que combina gráficos con muchos tipos de bases de datos. El GIS también puede mostrar información espacial precisa y real con gráficos y textos de acuerdo con las necesidades reales, y puede integrar ubicaciones geográficas y atributos de datos correlacionados como un todo orgánico. (p.1)

El presente estudio se enfoca en la población de Chile, debido a que existe la información suficiente para llevar a cabo un análisis con datos consistentes. Al respecto, Vidal, Hoffmeister y Biagini (2013) detallaron:

La calidad de los registros de defunciones tiene una calidad aceptable, porque su exhaustividad es del 100% de las muertes, el registro de la edad es completo, con certificación médica, codificación y validación centralizadas a partir de estándares internacionales, y porque las proporciones de causas de muerte mal definidas y neoplasias sin localización son bajas. (p. 2)

Por otro lado, Cornejo (2017) señaló: “La información no se explota completamente con fines preventivos” (p. 2) y que “Esto es importante porque Chile está disponible con mucha información relacionada con la salud, así como con el concepto de mortalidad y todo lo que la rodea” (p. 2). Asimismo, los especialistas del Instituto Nacional de Estadística de Chile (2019) indicaron que el 69% del total de muertes ocurridas en 2016 se debieron a solo cuatro grupos principales de causas de muerte:

1. Enfermedades del sistema circulatorio (28,148 muertes)
2. Tumores malignos (26,027 muertes)

3. Enfermedades del sistema respiratorio (9,847 muertes)
4. Causas externas de morbilidad y mortalidad (7,773 muertes). (p. 13)

Frente a esta problemática de elemental importancia, con la pretensión de colaborar con la población de la República de Chile y otras localidades, en base a conocimientos para el análisis y predicción hacia el futuro, se planteó llevar a cabo un sistema geolocalizado de pronóstico de defunciones para la toma de decisiones con el fin de informar el alcance de la cantidad de personas que mueren en un espacio y tiempo definido dentro de cada región del país, con el fin de proporcionar posibles soluciones para reducir la tasa de mortalidad.

Este estudio se justifica de manera teórica porque a través de su aplicación se podrá encontrar patrones con el fin de proponer estrategias para reducir la mortalidad. También, este estudio se justifica de manera económica porque a través de su aplicación se podrá tomar decisiones sobre la inversión económica correcta para combatir las principales enfermedades que amenazan a cada región de Chile. En los siguientes párrafos se sustenta respectivamente las razones de las justificaciones descritas.

En los últimos años, el Knowledge Discovery in Databases (KDD) y las técnicas de Data Mining (DM) de manera más específica, se están volviendo cada vez más útiles para procesar y explotar datos médicos (Neto, Brito, Lopes, Peixoto, Abelha y Machado, 2019, p. 16). Además, Neto et al. (2019) indicaron que la información útil descubierta y los patrones obtenidos con la aplicación de estos métodos, analizando en tiempo real datos complejos, heterogéneos y sacando conclusiones al respecto pueden ser utilizados por profesionales de la salud para determinar diagnósticos, pronósticos y tratamientos para pacientes en organizaciones de salud. (p. 16)

Las técnicas de minería de datos ayudan a encontrar el conocimiento oculto en un grupo de datos de enfermedades que pueden usarse para analizar y predecir el comportamiento futuro de las enfermedades (Ahmed, 2017, p. 3). Asimismo, Taranu (2015) concluyó:

Hoy en día, se han realizado muchos esfuerzos con el objetivo de una aplicación exitosa de la minería de datos en las instituciones de salud. El potencial principal de esta técnica radica en la posibilidad de investigar patrones ocultos en conjuntos de datos en el dominio de la salud. Estos patrones pueden usarse para el diagnóstico clínico. Sin embargo, los datos médicos crudos disponibles están ampliamente distribuidos, son diferentes y voluminosos por naturaleza. Estos datos deben recopilarse y almacenarse en almacenes de datos en formas organizadas y pueden integrarse para formar el sistema de información del hospital. (p. 38)

En la industria de la salud, la minería de datos ofrece varios beneficios, tales como: la detección del fraude en el seguro de salud, la disponibilidad de soluciones médicas para los pacientes a un costo menor, la detección de causas de enfermedades y la identificación de métodos de tratamiento médico. (Alby y Shivakumar, 2016, p. 1) También, Alby y Shivakumar (2016) indicaron que la minería de datos ayuda a los investigadores sanitarios a elaborar políticas sanitarias eficientes, construir sistemas de recomendación de medicamentos, desarrollar perfiles de salud de las personas, etc (p. 24). Además, Yoo, Alafaireet, Marinov, Peña-Hernández, Gopidi Chang y Hua (2012) indicaron:

La aplicación exitosa de la minería de datos proporciona nuevos conocimientos biomédicos y de atención médica que se pueden utilizar de manera efectiva para apoyar la toma de decisiones clínicas tales como: el proceso de diagnóstico, la elección de opciones de tratamiento y la predicción del pronóstico, así como la toma de decisiones administrativas, tales como: estimaciones de la dotación de personal, seguros, tendencias demográficas y de mercado, garantía de calidad, eficiencia de procesos, etc. (p. 2)

Sobre la base de la realidad problemática presentada se planteó el problema general y los problemas específicos de la investigación. El problema general de la investigación fue: ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo? Los problemas específicos de la investigación fueron los siguientes:

- **PE1:** ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple?
- **PE2:** ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal?
- **PE3:** ¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de series de tiempo?

El objetivo general fue determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo. Los objetivos específicos fueron los siguientes:

- **OE1:** Determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple.
- **OE2:** Determinar la precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal.
- **OE3:** Determinar la precisión del pronóstico de mortalidad con la técnica de series de tiempo.

La hipótesis general de la presente investigación fue: “El uso de técnicas de minería de datos generó mejores resultados en términos de incremento de precisión y reducción del error en el pronóstico de mortalidad”. Las hipótesis específicas fueron las siguientes:

- **HE1:** La precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple fue de al menos 80%.

Esta hipótesis fue propuesta en base a diferentes investigaciones sobre el pronóstico de la mortalidad utilizando la técnica de regresión lineal múltiple. Al respecto, Lin y Tsai (2015) explicaron:

En las pruebas empíricas, los resultados de la regresión utilizando los datos de mortalidad para el Reino Unido, Estados Unidos y Japón muestran una bondad de ajuste satisfactoria, lo

que nos convence de la idoneidad del supuesto lineal. Las ilustraciones empíricas muestran además que el rendimiento de nuestro modelo de ajuste y pronóstico de las tasas de mortalidad es bastante satisfactorio en comparación con los modelos de mortalidad conocidos existentes. (p. 1)

La regresión lineal suele tener un buen desempeño con datos de la mortalidad a través del tiempo. En este sentido, Poudel, Huang, Neupane y Steel (2017) señalaron:

Este estudio retrospectivo fue el primer estudio en la historia de la investigación del cáncer en Nepal que incluyó datos de 10 años del registro nacional de cáncer en hospitales para predecir la incidencia nacional de cáncer. Como tanto la tasa de incidencia de cáncer como el año siguieron al lineal, se verificó que la regresión lineal simple era la mejor opción para la incidencia de cáncer y el año. (p. 167)

- **HE2:** La precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal fue de al menos 60%.

Esta hipótesis fue propuesta en base a diferentes investigaciones sobre el pronóstico de la mortalidad utilizando la técnica de regresión logística ordinal. Al respecto, Hsieh, Li-Min, Cheng-Li, Meng-Ju, Chung-Y y Chia-Hung Kao(2018) sustentaron:

Nuestro análisis indicó que el modelo LR en lugar del modelo ANN proporcionó un método más apropiado para predecir el cáncer de páncreas en pacientes con DM2 en Taiwán. Nuestros hallazgos pueden aumentar el pronóstico del cáncer de páncreas a través de la vigilancia, el diagnóstico temprano y el tratamiento en personas con ciertos factores de riesgo. Se requieren investigaciones adicionales de otros países para



determinar si nuestros hallazgos son aplicables en otros lugares.  
(p. 6)

En la investigación de los autores mencionados anteriormente obtuvieron una precisión del 72.7%, lo que indica que el modelo es bueno. Por otro lado, Iriondo, Thio, Ruth del Río, Baucells, Bosio y Figueras (2020) detallaron: “El uso de modelos dinámicos en la mortalidad individual puede mejorar las predicciones de resultados. Se necesitan estudios de validación que prueben nuestra calculadora en línea en diferentes entornos y poblaciones prospectivas para confirmar su precisión y generalización” (p.10). La investigación mencionada anteriormente trabajó con 3 modelos de los cuales salió el pseudo r cuadrado de Nagelkerke: 0.329, 0.405 y 0.448 respectivamente.

- **HE3:** El error porcentual absoluto medio del pronóstico de mortalidad con la técnica de series de tiempo fue inferior al 5%.

Esta hipótesis fue propuesta en base a diferentes investigaciones sobre el pronóstico de la mortalidad utilizando la técnica de series de tiempo con el modelo ARIMA. Al respecto, Lin, Min, Guowei, Wu y Lin (2015) argumentaron:

El gobierno necesita con urgencia evaluar la pérdida causada por las lesiones, utilizando métodos estadísticos como series de tiempo. Nuestro enfoque de modelado muestra que es factible aplicar modelos de series temporales ARIMA para pronosticar la mortalidad por lesiones en Xiamen. Los modelos ARIMA basados en datos de vigilancia históricos son herramientas importantes para monitorear y pronosticar lesiones. (p. 7)

El modelo ARIMA es útil para el pronóstico de enfermedades debido a que se puede adaptar a cortos períodos de tiempo y se puede aplicar en cualquier nivel geográfico del País. Teniendo un pronóstico preciso y un análisis detallado. Por eso, Nayak y Narayan (2019) indicaron:

Se seleccionó ARIMA estacional  $(1,0,0) (0,1,1)_{12}$  como el modelo más apropiado para predecir la incidencia futura de casos de fiebre del dengue en el próximo año, lo cual es útil para los administradores de atención médica para una mejor preparación. El modelo puede simplificarse para incluir datos actuales y corregir el modelo. Se podrían desarrollar modelos predictivos más complejos teniendo en cuenta la precipitación, el período de incubación extrínseca, etc., para una predicción más precisa. El modelo se puede aplicar durante períodos de tiempo más pequeños o en niveles geográficos, como el nivel de distrito, para predecir la incidencia de la fiebre del dengue. (p. 5)

Por lo mencionado anteriormente, se puede ver que los modelos ARIMA son buenos para el pronóstico de la mortalidad. En consecuencia, Xie (2017) indicó:

Al utilizar el modelo ARIMA y los scripts SARIMA para 160 observaciones de datos, el autor encuentra algunos puntos en las series de tiempo. Los resultados indican que las tendencias en las tasas de incidencia de cáncer para 2015-2020 han disminuido en 300-550 por cada 100,000 habitantes en los Estados Unidos cada año. (p. 1)

La forma más utilizada para medir el rendimiento de las series temporales es a través de métricas de error. Si es que tienen un valor bajo se podría decir que es un buen modelo. Al respecto, Köppelová y Svatošová (2019) indicaron:

El valor del porcentaje medio absoluto de error (MAPE) no excedió el 5% en la mayoría de los modelos, lo que respalda la alta calidad de los modelos. Mientras el valor de la medida MAPE fluctúe dentro de un intervalo de no más de hasta el 10%, el modelo dado aún puede tomarse como uno de calidad relativamente buena. (p. 47)

En series de tiempo la precisión del modelo se mide a través de los errores de los cuales el MAPE es el más utilizado. Al respecto, Köppelová y Jindrová (2019) detallaron:

Se han alcanzado valores de MAPE inferiores al 15%, lo que expresa la calidad de los modelos para su uso en el análisis de series temporales estudiadas. Los valores métricos de MAPE se movieron con mayor frecuencia dentro del intervalo de 5 a 10% o alcanzaron valores inferiores al 5%. No existe un límite generalmente aceptado para la medida MAPE. En la práctica, es posible cumplir con diversas situaciones y el valor MAPE (su altura) se mueve solo dependiendo de la situación real. Es posible cumplir con una situación en la que se requiere un valor del 5%, en otro caso nuevamente del 15%, sin embargo, un modelo con un valor MAPE de aproximadamente el 10% puede tomarse como el adecuado. (p. 78)

## **II. MARCO TEÓRICO**

En esta segunda parte se detallará los trabajos relacionados con la investigación donde algunos fueron: predicción de la mortalidad hospitalaria temprana, pronóstico de la mortalidad en las provincias canadienses , etc. Pero no se encontraron investigaciones que tengan el mismo objetivo. También se explican las teorías relacionadas, como por ejemplo la importancia de la minería de datos. Se detalló las principales metodologías de minería de datos, esto se dio mediante un cuadro comparativo de puntaje ,donde se optó por usar la metodología CRISP-DM que contiene secciones o capas que ayudarán a administrar mejor el tiempo y la manera de predicción. Por último también se describe las diferentes técnicas de minería de datos.

Awad, Bader-EI-Den, McNicholas y Briggs (2017) investigaron sobre la predicción de la mortalidad hospitalaria temprana de pacientes en unidades de cuidados intensivos, donde la muestra que se tomo fue desde del rango de edad 16 años a más. También, explicaron que las tecnicas utilizadas fueron Random Forest (RF), árboles de decisión predictiva (DT), probabilísticos Naive Baye (NB) y modelos basados en reglas de teorías de resonancia adaptativa proyectiva (PARTE). (p. 1) De igual manera, Awad et al.(2017) explicaron que el marco propuesto excedió al puntaje estándar métodos SOFA, SAPS-I, APACHE-II, NEWS y Qsofa en términos de AUROC y tiempo, lo que significa que solo se usaron 6 horas en comparación con 48 horas o más después de la aceptación. (p. 2)

Li, Shrestha y Hu (2017) analizaron los accidentes fatales de tráfico vial (FARS) utilizando técnicas de minería de datos que tiene como objetivo aplicar análisis estadísticos y minería al grupo de datos del año 2007 informados por la Administración Nacional de Seguridad y Transportes en Carreteras. (p. 1) Al respecto, Li et al.(2017) concluyeron que la mortalidad por accidentes fatales gracias a métodos, metodologías y técnicas que se deben más a factores humanos que a factores ambientales, condiciones de superficie, clima o luz en el área. Donde el resultado se trata de personas intoxicadas o en estado de ebriedad. (p. 5)

Bergeron-Boucher, Simonacci, Oeppen y Gallo (2018) investigaron el pronóstico de la mortalidad en las provincias canadienses, el método propuesto

en esta investigación es identificar y analizar el patrón de edad específico de la provincia y el índice de tiempo, pero también identificar las tendencias de edad y tiempo comunes a todas las provincias. (p. 1) Al respecto, Bergeron-Boucher et al.(2018) explicaron que los resultados revelan que las provincias canadienses tienen un patrón de mortalidad muy equitativo, en la recomendación usan tendencias comunes de edad y tiempo considerando los marcadores específicos de la provincia (p. 6).

Islam, Hasan, Wang, Germack y Noor-E-Alam (2018) investigaron la aplicación y la perspectiva teórica de la minería de datos en el análisis de la atención médica, cuyo objetivo es presentar una revisión de la literatura sobre análisis de salud utilizando minería de datos y big data (p. 1). Asimismo, Islam et al. (2018) explicaron que se realizó un estudio de elementos de informe preferido para revisiones sistemáticas y metanálisis (PRISMA), también se realizó una búsqueda en la base de datos entre 2005 y 2016, elementos críticos de los estudios seleccionados: áreas de atención médica, técnicas de extracción de datos, tipo de análisis y fuente de datos. (p. 9)

Gaeta, Campanella, Gentile, Schifino, Capasso, Bandera, Banf, Arpesella y Ricci (2017) investigaron la mortalidad cardiovascular europea que tiene como objetivo proporcionar un informe exhaustivo para representar las muertes por enfermedades circulatorias en Europa en los últimos 30 años (p. 1). De igual manera, Gaeta et al. (2017) argumentaron que el estudio se basó en el sistema de información estadística de la OMS, base de datos de mortalidad y los 28 países pertenecientes a la Unión Europea, donde se consideran y evalúan los atributos de sexo y edad del conjunto de datos de todas las mujeres, muertes por tipos de enfermedades circulatorias, cardíacas, isquémicas, cerebrovasculares, pulmonares y otras enfermedades circulatorias de lo que se realizó en 2016 (p. 4).

Sharma, Narayan y Khatri (2016) aplicaron la tecnología de minería de datos para diagnosticar enfermedades, predecir y comprender profundamente los datos médicos. De igual manera, Sharma et al. (2016) explicaron que la minería de datos a lo largo del tiempo ha evolucionado más rápido en beneficio

de la tecnología, esto está contribuyendo a la biomedicina y la ciencia. (p. 1) Asimismo, Sharma et al.(2016) explicaron que el estudio se basó en enfermedades mortales, por eso se propuso un enfoque innovador para reducir esta tasa de falsa alarma que ayudaría al diagnóstico temprano de la enfermedad (p. 2).

Huerta, Huesca, Cabrera, Ruiz y Hernández (2015) investigaron el proceso de minería de datos para la clasificación de las causas de muerte fetal (p. 1). Al respecto, Huerta et al.(2015) explicaron que realizaron pruebas sobre el tipo de algoritmos de clasificación, donde la técnica de árboles presentó el mejor rendimiento en la comparación que se realizó, asimismo, se realizaron algoritmos para la selección de atributos en la base de datos, donde como resultado salieron 13 a 48 atributos del repositorio de datos sin interferir en la clasificación del estudio. (p. 5)

En esta sección se detalla las definiciones, modelos, técnicas, algoritmos, metodologías de minería de datos y herramientas de inteligencia de negocio. A partir de las cuales se pueden hacer comparaciones y análisis para apoyar la investigación. Además, esta sección permitirá tener mejores conceptos para una mejor comprensión de la investigación.

Uno de los conceptos clave que tiene la investigación es el uso de la minería de datos porque ayudará a encontrar patrones dentro de una gran cantidad de datos con el fin de generar conocimiento útil que sirva como estrategia y valor agregado. Al respecto, Kumar y Kaur (2016) indicaron:

La minería de datos (a veces llamada conocimiento o descubrimiento de datos) es el proceso de analizar datos desde diferentes perspectivas y resumirlos en información útil, información que puede usarse para aumentar los ingresos, reducir costos o ambos. El software de minería de datos es una de varias herramientas analíticas. Permite a los usuarios analizar datos desde muchas dimensiones o ángulos diferentes, clasificarlos y resumir las relaciones identificadas. Técnicamente, la

minería de datos es el proceso de encontrar correlaciones o patrones entre docenas de campos en grandes bases de datos relacionales. (p. 1)

La minería de datos ayuda a las empresas a depurar datos repetitivos para poder usar la información necesitada y así optimizar el trabajo. Al respecto, Hopping, Hellard, Millman (2018) detallaron:

La minería de datos permite a las empresas analizar todo el ruido caótico y repetitivo en sus datos y comprender lo que es relevante, luego hacer un buen uso de esa información para evaluar los resultados probables. El proceso identifica patrones e ideas que no se pueden encontrar en otros lugares, y al usar procesos automatizados para encontrar información específica, no solo acelera el tiempo que lleva encontrar los datos, sino que también aumenta la confiabilidad de los datos. (p. 1)

Las metodologías de minería de datos son fundamentales debido a que son una guía de como aplicar correctamente el proceso al contexto. Una de las metodologías mas usadas es CRISP-DM (*Cross Industry Standard Process for Data Mining*). Al respecto, Wirth y Hipp (2000) explicaron:

CRISP-DM es un modelo de proceso que proporciona un marco para llevar a cabo proyectos de minería de datos que es independiente de la industria, sector y la tecnología utilizada. El modelo de proceso CRISP-DM tiene como objetivo hacer que los grandes proyectos de minería de datos sean menos costosos, más confiables, más repetibles, más manejables y más rápidos. (p. 1)

Por otro lado, los especialistas de IBM (s. f.) indicaron: “Un esquema dimensional separa físicamente las medidas que cuantifican el negocio de los elementos descriptivos (también llamados dimensiones) que describen y categorizan el negocio” (p. 1). Estos esquemas pueden ser: modelo estrella, modelo copo de nieve y modelo constelación. Para la presente investigación, se utilizó una herramienta de geolocalización para la proyección de mapas. Asimismo, Longley, Goodchild, Maguire y Rhind (2005) indicaron:



GIS es mucho más que un contenedor de mapas en forma digital. Esta puede ser una descripción engañosa, pero es una definición útil para dar a alguien que busca una explicación simple: un cóctel invitado, familiar o vecino del asiento en un vuelo de una aerolínea. Todos conocemos y apreciamos el valor de los mapas, y la noción de que los mapas podrían ser procesados por una computadora es claramente análoga al uso de procesadores de texto u hojas de cálculo para manejar otros tipos de información. Un SIG es una herramienta para revelar lo que de otro modo sería invisible en la información geográfica. Una definición interesante que enfatiza el poder de un SIG como motor de análisis, para examinar datos y revelar sus patrones, relaciones y anomalías, cosas que podrían no ser obvias para alguien que mira un mapa. Un SIG es una herramienta para realizar operaciones en datos geográficos que son demasiado tediosos o caros o inexactos si se realizan a mano, una definición que habla de los problemas asociados con el análisis manual de mapas, particularmente la extracción de mediciones simples, por ejemplo. Todos tienen su propia definición favorita de un SIG, y hay mucho para elegir. (p. 16)

El pronóstico es la palabra clave de la presente investigación, debido a que es lo que se quiere obtener a través de un modelo obtenido con las técnicas de minería de datos con el fin de visualizar eventos futuros. Por lo mismo, Montgomery, Jennings y Kulahci (2015) indicaron:

Un pronóstico es una predicción de algún evento o eventos futuros. Por lo tanto, es una contribución crítica en muchos tipos de procesos de planificación y toma de decisiones. Un proceso es una serie de actividades conectadas que transforman una o más entradas en una o más salidas. Todas las actividades laborales se llevan a cabo en procesos, y el pronóstico no es una excepción. Las actividades en el proceso de pronóstico son: (a) Definición del problema, (b) Recopilación de datos, (c) Análisis de datos, (d) Selección y ajuste del modelo, (e) Validación del modelo, (f) Despliegue del modelo de pronóstico y (g) Monitorear el desempeño del modelo de pronóstico. (p. 12)

La precisión y el error porcentual absoluto medio (MAPE) serán las dimensiones a evaluar en la presente investigación. Por esto, Powers (2007) explicó:

La precisión o la confianza (como se llama en Data Mining) denota la proporción de casos positivos pronosticados que son correctamente reales positivos. Esto es en lo que se enfocan el aprendizaje automático, la minería de datos y la recuperación de información, pero se ignora por completo en el análisis ROC. Sin embargo, de manera análoga se le puede llamar verdadera precisión positiva (tpa), ya que es una medida de precisión de los positivos pronosticados en contraste con la tasa de descubrimiento de los positivos reales. (p.2)

Por otro lado, Basson, Kilbourn y Walters (2019) indicaron: “El MAPE se usó como la métrica principal para la evaluación de la precisión del pronóstico, ya que es ideal para evaluar grandes volúmenes de datos debido a su sensibilidad de escala y la facilidad de comparación asociada del mismo” (p. 5). De manera similar, Bowers y Jarrell (1992) determinaron:

El MAPE mide el tamaño típico de un error relativo al tamaño del valor que se está ejecutando. Aquí el error siempre se expresa en puntos porcentuales, lo que facilita la comparación de precisión entre el pronóstico de varias variables diferentes y entre varios pronósticos de la misma variable. En MAPE, un error relativo, la relación del valor absoluto del error con el valor real, se calcula primero para cada período. Entonces, el promedio de estas razones se expresa como un porcentaje. (p. 10)

La investigación aborda el pronóstico de mortalidad, por lo que es importante definir el concepto de mortalidad. Por consiguiente, los especialistas de la OMS (s. f.) explicaron:

Los datos de mortalidad indican el número de muertes por ubicación, intervalo de tiempo y causa. Datos de mortalidad de la OMS, muertes recopiladas en los sistemas nacionales de registro civil, con las causas

básicas de muerte codificadas por las autoridades nacionales. La causa básica de muerte se define como "la enfermedad o lesión que desencadenó la sucesión de eventos patológicos que condujeron directamente a la muerte, o las circunstancias del accidente o acto de violencia que produjo la lesión fatal", como se establece en la Clasificación Internacional de Enfermedades. (párr. 1)

Por otro lado, la Clasificación Internacional de Enfermedades es importante para establecer un orden estándar y de fácil comprensión. Por lo tanto, los especialistas de la OMS (2019) indicaron:

CIE es la base para la identificación de tendencias y estadísticas de salud a nivel mundial, y el estándar internacional para informar enfermedades y afecciones de salud. Es el estándar de clasificación de diagnóstico para todos los propósitos clínicos y de investigación. La CIE define el universo de enfermedades, trastornos, lesiones y otras condiciones de salud relacionadas, enumeradas de manera comprensiva y jerárquica que permite: (a) fácil almacenamiento, recuperación y análisis de información de salud para la toma de decisiones basada en evidencia, (b) compartir y comparar información de salud entre hospitales, regiones, entornos y países y (c) comparaciones de datos en la misma ubicación en diferentes períodos de tiempo. (párr. 1)

### **III. MÉTODO**

En esta sección, el método que se utilizó para llevar a cabo el estudio del sistema de pronóstico geocalizado de defunciones, en donde el estudio fue no experimental con diseño de investigación longitudinal de tendencia porque varía la cantidad de defunciones conforme los años, la variable fue el efecto de un sistema de pronóstico geocalizado que aplica la minería de datos para pronosticar la mortalidad. La población se extrajo del Departamento de Estadística e Información de Salud (DEIS) del gobierno de Chile, de 1997 a 2017, correspondiente a un aproximado de 1,900,000 registros de defunciones. La técnica que se utilizó para tener una precisión fiable fue la regresión lineal múltiple. La regresión ordinal simple fue usada para arrojar las probabilidades de muerte respecto a los indicadores. Las series temporales, en este caso el modelo Arima dio un pronóstico de la cantidad de muertes del año 2018.

### **3.1 Tipo y diseño de investigación**

El tipo de estudio fue no experimental porque las relaciones entre las variables se estudian sin manipulación de variables, es decir, se pueden estudiar múltiples variables. Al respecto, Hernández, Collado y Baptista (2014) explicaron: "Es el tipo de estudio que se realiza sin manipulación variable deliberada y donde solo se observan los fenómenos en su entorno natural para analizarlos" (p. 152).

El diseño de la investigación fue no experimental, longitudinal y de tendencias, ya que la investigación seleccionó datos de las defunciones 1997-2016, considerada adecuada para el análisis y desarrollo de un nuevo modelo de pronóstico de la información seleccionada. Asimismo, Hernández, Collado y Baptista (2014) indicaron: "Analiza los cambios a lo largo del tiempo en categorías, conceptos, variables o sus relaciones de alguna población general. Su característica distintiva es que el foco está en la población o el universo" (p. 160).

### **3.2 Variables y operacionalización**

**Variable:** Efecto de un sistema de pronóstico geocalizado que aplica técnicas de minería de datos para pronosticar la mortalidad.

**Definición conceptual:** Un sistema informático es un conjunto de recursos tecnológicos, humanos y económicos, interrelacionados entre sí, para satisfacer las necesidades de información de una institución, para una gestión y toma de decisiones adecuadas (Enríquez, 2016, p. 1). En este proyecto de investigación el sistema informático se basa en:

- Sistema geolocalizado: Es una herramienta para revelar lo que de otro modo sería invisible en la información geográfica. Una definición interesante que enfatiza el poder de un GIS como motor de análisis, para examinar datos y revelar sus patrones, relaciones y anomalías, cosas que podrían no ser obvias para alguien que mira un mapa. (Longley, Goodchild, Maguire y Rhind, 2005, p. 16)
- Minería de datos: La minería de datos ofrece numerosas formas de descubrir patrones ocultos en grandes cantidades de datos. Estos patrones ocultos pueden usarse potencialmente para predecir el comportamiento futuro. Los buenos datos son el primer requisito para una buena exploración de datos. Existen varias técnicas y algoritmos que se pueden usar para realizar la extracción de datos, pero su uso depende de la aplicación. Las técnicas de minería de datos predictivas son apropiadas cuando tiene un valor objetivo específico que desea predecir sobre sus datos. (Lico, 2017, p. 49)

**Definición operativa:** El uso de grandes bases de datos para encontrar y generar patrones repetitivos que ayuden a obtener información adecuada cuyas métricas se expresan en funcionalidad (nivel de precisión del pronóstico) y eficiencia (reducción de MAPE).

### 3.3 Población, muestra y muestreo

La población es el conjunto finito o infinito de elementos con características comunes (Arias, 2012, p. 81). La población de la investigación está compuesta por información sobre personas fallecidas que se recopila en la base de datos del Departamento de Estadística e Información de Salud (DEIS) del gobierno de

Chile, de 1997 a 2017, correspondiente a un aproximado de 1, 900,000 registros. A continuación, se especificarán los criterios de inclusión y exclusión.

**Criterio de inclusión:**

- Registros de personas fallecidas que tienen una codificación correcta de enfermedades según la CIE-10.
- Registros de personas fallecidas que tienen información sobre el lugar de residencia.

**Criterio de exclusión:**

- Registros de personas fallecidas que no tienen la codificación de la CIE-10. Por esta razón, los datos obtenidos del año 1990 a 1996 se eliminarán debido a que la codificación no pertenecía a ningún estándar internacional de salud.
- Registros de personas fallecidas que no tienen información relevante.

### **3.4 Técnicas e instrumentos de recolección de datos**

#### **Datos secundarios**

Los datos secundarios se refieren a la información que no se ha producido explícitamente para los fines de la investigación, sino que preexiste y el investigador la hace rentable (González, 1997, p. 244).

#### **Validez**

La validez es la relevancia de un instrumento de medición, se refiere a la precisión con la que el instrumento mide lo que pretende medir, es decir, es la metodología de un instrumento para representar o describir el atributo de interés para el examinador (Ñaupas et al., 2014, p. 215).

En cuanto a la validez del contenido, Naúpas et al. (2014) indicaron que se refiere al "grado en que un instrumento refleja un dominio o contenido en particular, y para que haya validez de contenido, el instrumento debe estar relacionado con los objetivos, competencias y contenido del curso" (p. 215). Además, Naúpas et al. (2014) mencionaron: "La validez de constructo se refiere al grado de correspondencia entre los resultados de una prueba y los conceptos teóricos en los que se basan los sujetos a medir" (p. 216).

## **Fiabilidad**

Es un instrumento de medición que se refiere al grado en que su aplicación repetitiva al mismo objeto o individuo produce resultados iguales, veraces y consistentes basados en su validez y cumplimiento de objetivos, por ejemplo, una prueba de inteligencia (Cociente de inteligencia, IQ) se aplica a un grupo de personas que obtienen la inteligencia actual, después de un mes se aplica la misma prueba, pero se proporcionan valores diferentes, ya que en mediciones posteriores, dicha prueba no sería confiable (Hernández, Fernández y Baptista, 2010). Para esta investigación, el registro de observación utilizado como instrumento no requiere cálculos de confiabilidad, ya que la información se extrae directamente de la base de datos de la Dirección de Estadísticas de Información de Salud (DEIS).

## **3.5 Procedimientos**

La recopilación de datos se realizó utilizando una base de datos proporcionada por el país de Chile, a través de su institución llamada Departamento de Estadística e Información de Salud, donde se comparte la información para las investigaciones pertinentes. La base de datos de Defunciones de 1997 a 2017 fue elegida debido a que las causas de muerte estaban codificadas bajo el estándar internacional de la OMS (CIE-10).



### 3.6 Método de análisis de datos

Debido a que la investigación se basa en el análisis de los datos a través de técnicas de minería de datos, se usó 3 técnicas de los cuales se irán explicando a continuación:

- **Regresión lineal múltiple:** Esta técnica evalúa cómo se relacionan las variables. Para hacer esto, se agregan poderes a la ecuación de regresión y se observa si estos poderes mejoran significativamente la predicción (Pérez y Medrano, 2010). Esta técnica se usó para ver la importancia de las variables en los pronósticos de muertes, también se usó para predecir el número de muertes para el año siguiente. Asimismo, Cheng, Wai-Ki, Hou y Aoki-Kinoshita (2015) indicaron: “El modelo MLR es un método tradicional que intenta modificar la relación entre las variables independientes  $x_{ij}$ , ( $j = 1, 2, 3, n$ ) con variables independientes  $y_i$ , ( $i = 1, 2, 3, k$ )” (p. 152).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in} + \epsilon_i, \quad i = 1, 2, \dots, k$$

or equivalently

$$\mathbf{y} = \mathbf{x}\beta + \epsilon.$$

*Figura 1: Ecuación de RLM, adaptado de "A Multiple Linear Regression Model for structure of N-Linked oligosaccharides", por Cheng et al., 2015, the institution of Engineering & Technology*

- **Regresión logística ordinal:** El análisis de regresión logística es una técnica estadística multivariable para el análisis de una relación de dependencia entre una variable dependiente y un conjunto de variables independientes, de manera similar al funcionamiento del análisis de regresión lineal clásico. El objetivo del análisis es poder hacer predicciones de comportamiento, es decir, estimar las probabilidades de un evento definido por la variable dependiente en función de un conjunto de variables predictoras o pronosticadoras. (López-Roldan y Facheli, 2015) La técnica de regresión logística ordinal tenía el propósito de tener

probabilidades para cada variable, es decir, saber qué variable tiene más influencia en la variable independiente.

$$\text{Logit}[P(Y \leq j | X)] = \alpha_j - \sum_{K=1}^K \beta_K X_K, j = 1, 2, \dots, j-1$$

*Figura 2: Ecuación de RLO, adaptado de "A Multiple Linear Regression Model for structure of N-Linked oligosaccharides", por Cheng et al., 2015, The institution of Engineering & Technology*

- **Series de tiempo:** Una serie de tiempo es una colección de observaciones realizadas secuencialmente a lo largo del tiempo, como lo indican Angadi y Kulkarni (2015), quienes explicaron que las técnicas de análisis de datos de series de tiempo requieren información verificable como premisa para evaluar resultados futuros. Los datos de series de tiempo pueden definirse como datos numéricos recopilados en una secuencia particular durante un período de tiempo a intervalos regulares. Los datos de la serie temporal pueden incluir los valores recopilados al final de cada semana, mes, trimestre o año, etc. La intención es determinar si existe un vínculo entre los datos recopilados hasta el momento y cómo cambian los datos. (p. 104)

Las series de tiempo son modelos paramétricos que intentan obtener la representación de la serie en términos de la interrelación temporal de sus elementos. Yule y Slutsky propusieron este tipo de modelo que caracteriza las series como sumas o diferencias, ponderadas o no, de variables aleatorias o de las series resultantes. Fueron la base de los procesos de promedios móviles y regresiones de autores que han tenido un desarrollo espectacular después de la publicación en 1970 del libro Box-Jenkins sobre los modelos ARIMA. (Gonzales, 2009) Para los parámetros del modelo, "p" es el orden del proceso autorregresivo AR (p), "d" el orden de diferenciación o grado de integración del proceso I (d) y "q" el orden de la media móvil MA (q). En general, el modelo ARIMA (p, d, q) sigue la fórmula a continuación:

$$y_t = \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \mu + \varepsilon_t$$

Figura 3: Ecuación del modelo ARIMA, adaptado de “Temporal analysis of road accidents by ARIMA model”, por Gherida et al., 2018, International Journal of Innovation and Applied Studies.

### 3.7 Aspectos éticos

El estudio utilizó una base de datos de defunciones de acceso público del DEIS de Chile. Por lo que la información mostrada es genuina y veraz. El tratamiento y confidencialidad de los datos obtenidos es respaldada por la ley N° 19,628 que dice: “Toda persona puede efectuar el tratamiento de datos personales, siempre que lo haga de manera concordante con esta ley y para finalidades permitidas por el ordenamiento jurídico” (Biblioteca del Congreso Nacional de Chile, 2020, párr. 2).

La investigación respetó la autoría intelectual de las fuentes, citando y refiriéndose a ellas según las normas establecidas por la Universidad César Vallejo. De acuerdo al artículo 15 del código de ética en investigación. La universidad promueve la originalidad de las investigaciones y para ello tiene a disposición de la comunidad universitaria que investiga, el programa de evaluación de trabajos de investigación – Turnitin. Además, el artículo 44 del código de ética del Colegio de Ingenieros del Perú referidos a la divulgación de información o la omisión de autor o coautores que intervienen en la investigación las cuales se están cumpliendo.

## **IV. RESULTADOS**

En este capítulo se realizó la descripción detallada de los resultados obtenidos dentro de la investigación basándose en indicadores del “incremento de precisión” y “reducción del MAPE”. Se logró determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo. Los datos fueron analizados por cada región, es decir, las técnicas se aplicaron a las 16 regiones en las que se divide Chile. El mejor pronóstico para todas las regiones se muestra a continuación.

#### IV.1 Prueba de la hipótesis específica 1

A continuación, se detallará los resultados obtenidos con el modelo de regresión lineal múltiple. Donde la variable dependiente fue la cantidad de muertes y como variables independientes se usaron el año, mes, local de defunción, sexo y los conjuntos de enfermedades estipuladas en la CIE-10. Se llegó a medir el incremento de precisión al finalizar el uso del modelo.

En la tabla 1 se muestra la prueba ANOVA que indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 1: ANOVA del modelo de regresión lineal múltiple

Model	Sum of squares	Square Mean	F	Sig.
Regression	5537222.496	276861.125	3213.164	.000
Residue	6732670.752	6732670.752	-	-
Total	6732670.752	-	-	-

En la tabla 2 se puede ver que el modelo generado para la región de Valparaíso presenta un R cuadrado ajustado de 0.822, es decir que la varianza de la mortalidad se explica al 82.2% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados y el modelo es bueno.

Tabla 2: Resumen del modelo de regresión lineal múltiple

Model	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Valparaíso	0.822	0.822	9.28249	1.943

En la tabla 3 se muestra los estadísticos descriptivos correspondientes a la variable dependiente e independientes.

Tabla 3: *Estadístico descriptivo del modelo de regresión lineal múltiple*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	16.1329	22.01304
TYear	2007.0845	6.06591
TMonth	6.5177	3.44305
DLocationID	.5334	.49890
SexID	.5130	.49985
ID_A00_B99	.0689	.25325
ID_C00_D48	.0725	.25940
ID_E00_E90	.0725	.25940
ID_F00_F99	.0694	.25423
ID_G00_G99	.0718	.25809
ID_I00_I99	.0725	.25940
ID_J00_J99	.0725	.25940
ID_K00_K93	.0725	.25940
ID_L00_L99	.0419	.20033
ID_M00_M99	.0429	.20262
ID_N00_N99	.0723	.25904
ID_O00_O99	.0065	.08022
ID_P00_P96	.0369	.18857
ID_Q00_Q99	.0492	.21620
ID_R00_R99	.0617	.24071
ID_S00_T98	.0722	.25880

En la tabla 4 se indica los coeficientes no estandarizados (B) que ayudarán a formar la ecuación para la predicción y los coeficientes estandarizados ( $\beta$ ) indicarán la importancia de cada variable.

Tabla 4: *Coeficientes del modelo de regresión lineal múltiple*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-520.885	26.075	-	-19.976	.000	-
TYear	.261	.013	.072	20.122	.000	1.001
TMonth	.110	.023	.017	4.818	.000	1.000
DLocationID	-3.057	.161	-.069	-18.975	.000	1.042
SexID	-1.446	.158	-.033	-9.127	.000	1.011
ID_A00_B99	3.064	.482	.035	6.352	.000	2.406
ID_C00_D48	53.949	.478	.636	112.929	.000	2.476
ID_E00_E90	8.070	.478	.095	16.892	.000	2.476
ID_F00_F99	3.325	.482	.038	6.902	.000	2.419
ID_G00_G99	4.542	.479	.053	9.487	.000	2.461

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_I00_I99	65.998	.478	.778	138.152	.000	2.476
ID_J00_J99	19.818	.478	.234	41.484	.000	2.476
ID_K00_K93	13.152	.478	.155	27.531	.000	2.476
ID_L00_L99	-.278	.539	-.003	-.515	.606	1.880
ID_M00_M99	.309	.536	.003	.576	.565	1.900
ID_N00_N99	5.896	.478	.069	12.335	.000	2.472
ID_O00_O99	.479	1.051	.002	.455	.649	1.147
ID_P00_P96	2.259	.559	.019	4.040	.000	1.793
ID_Q00_Q99	1.455	.518	.014	2.807	.005	2.025
ID_R00_R99	3.457	.494	.038	6.999	.000	2.279
ID_S00_T98	12.588	.478	.148	26.325	.000	2.470

### Ecuación del modelo de regresión lineal múltiple

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Defunciones} = & - 520.88 + 0.26(\text{TYear}) + 0.11(\text{TMonth}) - 3.06 (\text{DLocationID}) - \\ & 1.45(\text{SexID}) + 3.06(\text{ID\_A00\_B99}) + 53.95(\text{ID\_C00\_D48}) + 8.07(\text{ID\_E00\_E90}) + \\ & 3.32(\text{ID\_F00\_F99}) + 4.54(\text{ID\_G00\_G99}) + 66(\text{ID\_I00\_I99}) + 19.82(\text{ID\_J00\_J99}) + \\ & 13.15(\text{ID\_K00\_K93}) + 0.31(\text{ID\_M00\_M99}) + 5.90(\text{ID\_N00\_N99}) + 0.48(\text{ID\_O00\_O99}) + \\ & 2.26(\text{ID\_P00\_P96}) + 1.45(\text{ID\_Q00\_Q99}) + 3.46(\text{ID\_R00\_R99}) + 12.59(\text{ID\_S00\_T98}) \end{aligned}$$

### Interpretación de los valores B

Los valores B indican la relación entre el número de muertes y cada predictor, si el valor es positivo, significa que existe una relación positiva entre el predictor y el resultado. Por otro lado, si el valor B es negativo, se concluye que tienen una relación negativa. Sin embargo, los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.261): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.261 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.110): Este valor indica que a medida que el mes aumenta en una unidad genera un crecimiento de 0.110 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -3.057): Este valor determina que cuando exista presencia en un hospital o clínica, la mortalidad cae en 3.057, manteniendo las otras condiciones constantes.

**SexID** (B = -1.446): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.446, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 3.064): Este valor indica que cuando exista presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 3.064, manteniendo las otras condiciones constantes.

**ID\_C00\_D48** (B= 53.949): Este valor indica que cuando exista presencia en enfermedades de tumores [neoplasias], la mortalidad aumenta a 53.949, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 8.070): Este valor indica que, cuando exista una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 8.070 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 3.325): Este valor determina que cuando exista presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 3.325, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 4.542): Este valor establece que, dado que hay una presencia en enfermedades del sistema nervioso central, generará un aumento de 4.542 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 65.998): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 65.998, manteniendo las otras condiciones constantes.



**ID\_J00\_J99** (B = 19.818): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 19.818 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_K00\_K93** (B = 13.152): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 13.152, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.309): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.309 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 5.896): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 5.896, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.479): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.479 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 2.259): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 2.259, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 1.455): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.455 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 3.457): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, genera un aumento de 3.457 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $B = 12.588$ ): Este valor establece que, como existe la presencia de lesiones, envenenamientos y algunas otras consecuencias de causas externas, generará un aumento de 12.588 en la mortalidad, manteniendo las otras condiciones constantes.

### **Interpretación de los valores $\beta$**

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.072$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.06), la mortalidad aumenta en 7.2% desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.585 ( $0.072 * 22.013$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.017$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.44), la mortalidad aumenta en un 1.7% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.374 ( $0.017 * 22.013$ ), Manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.069$ ): Este valor determina que cuando exista presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 6.9% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.519 ( $0.069 * 22.013$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.033$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en 3.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.726 ( $0.033 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.035$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.253), la mortalidad aumenta en un 3.5% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.770 ( $0.035 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.636$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.259), la mortalidad aumenta en un 63.6% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 14 ( $0.636 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.095$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.259), la mortalidad aumenta en un 9.5% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 2.091 ( $0.095 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.038$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.254), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.836 ( $0.038 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.053$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.258), la mortalidad aumenta en un 5.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.166 ( $0.053 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.778$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.259), la

mortalidad aumenta en un 77.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 17.126 ( $0.778 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.234$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.259), la mortalidad aumenta en un 23.4% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 5.151 ( $0.234 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.155$ ): Este valor indica que cuando hay una presencia de enfermedades del sistema digestivo con una desviación estándar (0.259), la mortalidad aumenta en un 15.5% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 3.412 ( $0.155 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.003$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.202), la mortalidad aumenta en un 0.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.06 ( $0.003 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.069$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.259), la mortalidad aumenta en un 6.9% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.518 ( $0.069 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.002$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.080), la mortalidad aumenta en un 0.2% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye

un cambio de 0.04 (0.002 x 22.013), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.019$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.188), la mortalidad aumenta en un 1.9% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.418 (0.019 x 22.013), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.014$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.216), la mortalidad aumenta en un 1.4% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.308 (0.014 x 22.013), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.038$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.240), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.836 (0.038 x 22.013), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.148$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.258), la mortalidad aumenta en un 14.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 3.257 (0.148 x 22.013), manteniendo las otras condiciones constantes.

### **Resumen de los modelos de regresión lineal múltiple**

El modelo fue aplicado a cada una de las 16 regiones de Chile. Todos los modelos estuvieron dentro del rango de 1.5 a 2.5 en la prueba del estadístico de Durbin Watson, lo que indica que los residuos no están autocorrelacionados.

En la tabla 5 se muestra los resultados obtenidos al aplicar regresión lineal múltiple en cada región de Chile ordenado de mayor a menor según el r cuadrado ajustado.

Tabla 5: *Resumen de los modelos de regresión lineal múltiple por región*

Región	R Square	Adjusted R Square	Durbin-Watson
Valparaíso	0.822	0.822	1.943
Metropolitana de Santiago	0.814	0.813	1.896
Biobío	0.787	0.787	1.903
Maule	0.748	0.748	1.944
Libertador General Bernardo O'Higgins	0.722	0.721	1.943
La Araucanía	0.702	0.701	1.871
Coquimbo	0.699	0.698	2.089
Ñuble	0.690	0.689	2.015
Antofagasta	0.679	0.678	2.022
Los Lagos	0.678	0.677	2.027
Los Ríos	0.670	0.669	1.967
Atacama	0.534	0.533	1.997
Magallanes y de la Antártica Chilena	0.531	0.530	2.035
Tarapacá	0.488	0.486	2.088
Arica y Parinacota	0.458	0.456	2.109
Aysén del Gral. C. Ibañez del Campo	0.280	0.277	1.944

## IV.2 Prueba de la hipótesis específica 2

A continuación, se detallará los resultados obtenidos con el modelo de regresión logística ordinal. Como variables se usaron el año, mes, local de defunción, sexo, rango de edad y los conjuntos de enfermedades estipuladas en la CIE-10. Se llegó a medir el incremento de precisión al finalizar el uso del modelo.

En la tabla 6 se muestra el Pseudo R2, donde se puede ver que Nagelkerke explica a un 66.9% el número de muertes.

Tabla 6: Resumen del modelo de regresión logística ordinal

Model	McFadden	Nalgerkerke	CoxSnell
Valparaíso	0.589	0.669	0.329

En la tabla 7 se muestra los coeficientes de cada variable, los valores del wald y el nivel de significancia de cada variable.

Tabla 7: Coeficientes del modelo de regresión logística ordinal

Variable	Coef	S.E.	Wald Z	Wald	Pr(> z )
y>=2	-162.0594	11.4281	-14.18	201.07	<0.0001
y>=3	-164.2689	11.4370	-14.36	206.21	<0.0001
y>=4	-167.2142	11.4504	-14.60	213.16	<0.0001
y>=5	-172.0126	11.4967	-14.96	223.80	<0.0001
TYear	0.0796	0.0057	13.99	195.72	<0.0001
TMonth	0.0394	0.0096	4.12	16.97	<0.0001
DLocationID	-1.6696	0.0749	-22.28	496.40	<0.0001
SexID	0.0107	0.0665	0.16	0.03	0.8718
RAgeID_0_10	-14.8020	46.5892	-0.32	0.10	0.7507
RAgeID_11_19	-15.8191	66.3770	-0.24	0.06	0.8116
RAgeID_20_64	-4.5121	0.1109	-40.70	1656.49	<0.0001
ID_A00_B99	-12.4566	61.6745	-0.20	0.04	0.8399
ID_C00_D48	3.9337	0.1450	27.14	736.58	<0.0001
ID_D50_D89	-12.8522	97.7578	-0.13	0.02	0.8954
ID_E00_E90	-4.4350	1.0087	-4.40	19.36	<0.0001
ID_F00_F99	-13.1575	79.0866	-0.17	0.03	0.8679
ID_G00_G99	-12.4642	56.0812	-0.22	0.05	0.8241
ID_I00_I99	5.1980	0.1527	34.04	1158.72	<0.0001
ID_J00_J99	0.8725	0.1548	5.63	31.70	<0.0001
ID_K00_K93	-3.0438	0.5171	-5.89	34.69	<0.0001
ID_L00_L99	-13.3223	114.7066	-0.12	0.01	0.9075
ID_M00_M99	-12.7434	92.9380	-0.14	0.02	0.8909
ID_N00_N99	-12.8375	66.9934	-0.19	0.04	0.8480
ID_O00_O99	-8.5912	303.1920	-0.03	0.001	0.9774
ID_P00_P96	-3.3835	105.2273	-0.05	0.003	0.9743
ID_Q00_Q99	-10.1618	77.4986	-0.13	0.02	0.8957
ID_R00_R99	-12.7384	66.0539	-0.19	0.04	0.8471

## Ecuación del modelo de regresión logística ordinal

Al reemplazar los coeficientes en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Defunciones} = & -b + 0.0796(\text{TYear}) + 0.0394(\text{TMonth}) - 1.6696(\text{DLocationID}) + \\ & 0.0107(\text{SexID}) - 14.802(\text{RAgeID}_0_{10}) - 15.819(\text{RAgeID}_{11}_{19}) - 4.512 \\ & (\text{RAgeID}_{20}_{64}) - 12.457(\text{ID}_{A00}_{B99}) - 3.9337 (\text{ID}_{C00}_{D48}) - 12.457 \\ & (\text{ID}_{D50}_{D89}) - 4.4350(\text{ID}_{E00}_{E90}) - 13.1575(\text{ID}_{F00}_{F99}) - 12.4642 \\ & (\text{ID}_{G00}_{G99}) + 5.198(\text{ID}_{I00}_{I99}) + 0.8725(\text{ID}_{J00}_{J99}) - 3.0438(\text{ID}_{K00}_{K93}) - 13. \\ & 3223(\text{ID}_{L00}_{L99}) - 12.743(\text{ID}_{M00}_{M99}) - 12.838(\text{ID}_{N00}_{N99}) - \\ & 8.591(\text{ID}_{O00}_{O99}) - 3.3835(\text{ID}_{P00}_{P96}) - 10.1618(\text{ID}_{Q00}_{Q99}) - \\ & 12.7384(\text{ID}_{R00}_{R99}) \end{aligned}$$

En la ecuación “b” puede tomar varios valores debido a que el número de defunciones se dividieron en 5 categorías para cada región, es por eso que el valor de “b” según los resultados puede tomar estos valores: -162.0594, -164.269, -167.214 y -172.0126.

## Interpretación de los valores B

Para la interpretación de los coeficientes de la regresión logística ordinal se deberá elevar E al coeficiente para medir la probabilidad.

Tabla 8: Probabilidades del modelo de regresión logística ordinal

Coefficients	B	Exp B
TYear	0.0796	1.082853839
TMonth	0.0394	1.040186475
DLocationID	-1.6696	0.18832238
SexID	0.0107	1.01075745



<b>Coefficients</b>	<b>B</b>	<b>Exp B</b>
RAgeID_0_10	-14.802	3.72883E-07
RAgeID_11_19	-15.819	1.3485E-07
RAgeID_20_64	-4.512	0.010975388
ID_A00_B99	-12.457	3.89195E-06
ID_C00_D48	3.9337	51.09568238
ID_D50_D89	-12.457	3.89195E-06
ID_E00_E90	-4.4350	0.011855066
ID_F00_F99	-13.1575	1.93095E-06
ID_G00_G99	-12.4642	3.86248E-06
ID_I00_I99	5.198	180.9100597
ID_J00_J99	0.8725	2.392885596
ID_K00_K93	-3.0438	0.047653462
ID_L00_L99	-13.3223	1.63757E-06
ID_M00_M99	-12.743	2.92154E-06
ID_N00_N99	-12.838	2.65916E-06
ID_O00_O99	-8.591	0.000185733
ID_P00_P96	-3.3835	0.033928497
ID_Q00_Q99	-10.1618	3.86177E-05
ID_R00_R99	-12.7384	2.93618E-06

**Tyear:** Por cada unidad en el año en defunciones, las probabilidades de ser más probable se multiplican 1.8 veces. Es decir, aumenta en un 0.0828 % (Significativo).

\* Por cada unidad en la disminución del año en defunciones, las probabilidades de que sea más probable se multiplican 1.08 veces. Es decir, decrece en un 0.0828 %.

**TMonth:** Por cada unidad en el mes en defunciones, las probabilidades de ser más probable se multiplican 1.04 veces. Es decir, aumenta en un 0.40%, (Significativo).

\* Por cada unidad en la disminución del mes en defunciones, las probabilidades de que sea más probable se multiplican 1.04 veces. Es decir, decrece en un 0.0401%.

**DLocationID:** Para las defunciones cuya localización de muerte ha sido en hospitales, las probabilidades de que se origine una defunción son 88.2 menores que no pertenecen a otra localización (Significativo).

\* Para las defunciones que no pertenecen a la localización de hospitales, la probabilidad de que fallezcan es de 8.47 veces a los que son de localización.

**SexID:** Para las muertes cuyo sexo ha sido femenino, la probabilidad de ser más probable es mayor es 1.01 veces. Es decir 10%, (Significativo).

\* Para las muertes cuyo sexo no era femenino (masculino), la probabilidad de ser menos probable es 1.01075745 veces mayor que la de las mujeres.

**RAgeID\_0\_10:** Para las muertes entre las edades de 0 a 10, la probabilidad de morir es 99.99996% menos que las muertes que no pertenecen al rango de edad de 0 a 10 (no significativo).

\* Para las muertes que no pertenecen al rango de edades de 0-10, la probabilidad de morir es 2681806.35749 veces en los rangos de edad de 0-10.

**RAgeID\_11\_20:** Para las muertes entre las edades de 11 a 19 años, la probabilidad de morir es 98.9% menor que las muertes que no pertenecen a las edades de 11 a 19 (no significativo).

\* Para las muertes que no pertenecen al rango de edad de 11 a 19 años, la probabilidad de morir es 91.07 veces en el rango de edad de 11 a 19.

**RAgeID\_21\_64:** Para las muertes entre el rango de edad de 20 a 64, la probabilidad de morir es 98.9% menor que las muertes que no pertenecen al rango de edad de 20 a 64.

\* Para las muertes que no pertenecen al rango de edad de 20 a 64, la probabilidad de morir es 91.07 veces en el rango de edad de 20 a 64, (Significativo).

**ID\_A00\_B99:** Para las muertes que pertenecen a enfermedades del grupo A00\_B99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo A00\_B99.

\* Para las muertes que no pertenecen al grupo A00\_B99, la probabilidad de que fallezca es de 256940.608178 veces para quienes pertenecen al grupo A00\_B99.

**ID\_C00\_D48:** Para las muertes que se deben a la enfermedad C00\_D48, la probabilidad de ser más probable es 51.09 veces. Es decir, 0,095% mayor que las que no pertenecen a la enfermedad C00\_D48, (Significativo).

\* Para las muertes que no se deben a la enfermedad de C00\_D48, la probabilidad de ser menos probable es 51.09 veces mayor que las causas que si pertenecieran a C00\_D48.

**ID\_D50\_D89:** Para las muertes que pertenecen a enfermedades del grupo D50\_D89, la probabilidad de que la muerte tenga un 99.9% menos que las muertes que no pertenecen al grupo D50\_D89.

\* Para las muertes que no pertenecen al grupo D50\_D89, la probabilidad de que fallezca es de 256940.608178 veces para quienes pertenecen al grupo D50\_D89.

**ID\_E00\_E90:** Para las muertes que pertenecen a enfermedades del grupo E00\_E90, la probabilidad de que la muerte sea 98.8% menor que las muertes que no pertenecen al grupo E00\_E90 (Significativo).

\* Para las muertes que no pertenecen al grupo E00\_E90, la probabilidad de que la atención sufra es 84.3521242311 veces mayor que la del grupo E00\_E90.

**ID\_F00\_F99:** Para las muertes pertenecientes a enfermedades del grupo F00\_F99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo F00\_F99.

\* Para las muertes que no pertenecen al grupo F00\_F99, la probabilidad de que fallezca es de 517879.800098 veces mayor que la del grupo F00\_F99.

**ID\_G00\_G99:** Para las muertes que pertenecen a enfermedades del grupo G00\_G99, la probabilidad de que la muerte tenga es 99.9% menor que las muertes que no pertenecen al grupo G00\_G99.

\* Para las muertes que no pertenecen al grupo G00\_G99, la probabilidad de que fallezca es de 258901.016963 veces mayor que la del grupo G00\_G9.

**ID\_I00\_I99:** Para las muertes que se deben a la enfermedad I00\_I99, la probabilidad de ser más probable es 180.91 veces. Es decir, 0.91% mayor que las que no pertenecen a la enfermedad I00\_I99, (Significativo).

\* Para las muertes que no se deben a la Enfermedad I00\_I99, la probabilidad de ser menos probable es 180.91 veces mayor que las causas que si pertenecieran a I00\_I99.

**ID\_J00\_J99:** Para las muertes que se deben a la enfermedad de J00\_J99, la probabilidad de ser más probable es 2.392885596 veces. Es decir, 0.39% más que las que no pertenecen a la enfermedad de J00\_J99, (Significativo).

\* Para las muertes que no se deben a la enfermedad J00\_J99, la probabilidad de ser menos probable es 2.392885596 veces mayor que las causas que si pertenecieran a J00\_J99.

**ID\_K00\_K93:** Para las muertes pertenecientes a enfermedades del grupo K00\_K93, la probabilidad de que la muerte sea 95.2% menor que las muertes que no pertenecen al grupo K00\_K93.

\* Para las muertes que no pertenecen al grupo K00\_K93, la probabilidad de que fallezca es de 20.89 veces mayor que las que están en el grupo K00\_K93.

**ID\_L00\_L99:** Para las muertes pertenecientes a enfermedades del grupo L00\_L99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo L00\_L99.

\* Para las muertes que no pertenecen al grupo L00\_L99, la probabilidad de que es de 610660.918312 veces mayor que la del grupo L00\_L99.

**ID\_M00\_M99:** Para las muertes pertenecientes a enfermedades del grupo M00\_M99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo M00\_M99.

\* Para las muertes que no pertenecen al grupo M00\_M99, la probabilidad de que fallezca es de 342285.23313 veces mayor que la del grupo M00\_M99.

**ID\_N00\_N99:** Para las muertes pertenecientes a enfermedades del grupo N00\_N99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo N00\_N99.

\* Para las muertes que no pertenecen al grupo N00\_N99, la probabilidad de que fallezca es de 376058.604973 veces la del grupo N00\_N99.

**ID\_O00\_O99:** Para las muertes pertenecientes a enfermedades del grupo O00\_O99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo O00\_O99.

\* Para las muertes que no pertenecen al grupo O00\_O99, la probabilidad de que fallezca es de 5384.07283574 veces la del grupo O00\_O99.

**ID\_P00\_P96:** Para las muertes pertenecientes a enfermedades del grupo P00\_P96, la probabilidad de que la muerte tenga es 99.9% menor que las muertes que no pertenecen al grupo P00\_P96.

\* Para las muertes que no pertenecen al grupo P00\_P96, la probabilidad de que fallezca es de 29.4737488666 veces mayor que la del grupo P00\_P96.

**ID\_Q00\_Q99:** Para las muertes que pertenecen a enfermedades del grupo Q00\_Q99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo Q00\_Q99.

\* Para las muertes que no pertenecen al grupo Q00\_Q99, la probabilidad de que fallezca es de 25894.8616826 veces mayor que la del grupo Q00\_Q99.

**ID\_R00\_R99:** Para las muertes pertenecientes a enfermedades del grupo R00\_R99, la probabilidad de que la muerte sea 99.9% menor que las muertes que no pertenecen al grupo R00\_R99.

\* Para las muertes que no pertenecen al grupo R00\_R99, la probabilidad de que fallezca es de 340599.455041 veces mayor que la del grupo R00\_R99.

## Resumen de los modelos de regresión logística ordinal

El modelo fue aplicado a cada una de las 16 regiones de Chile. En la tabla 9 se muestra los resultados obtenidos al aplicar regresión lineal múltiple en cada región de Chile ordenado de mayor a menor según el Pseudo R cuadrado de Nagelkerke.

Tabla 9: Resumen de los modelos de regresión logística ordinal por región

Región	Pseudo R		
	McFadden	Nalgerkerke	CoxSnell
Metropolitana de Santiago	0.6016495	0.6650979	0.2784912
Valparaíso	0.5899531	0.6695313	0.3295600
Biobío	0.4481945	0.575911	0.3717824
Maule	0.4730015	0.5763132	0.3261750
La Araucanía	0.562207	0.562207	0.364955
Libertador General Bernardo O Higgins	0.4415637	0.5506673	0.3231616
Los Lagos	0.4219182	0.5351612	0.3224431
Coquimbo	0.4696517	0.5901884	0.3681046
Ñuble	0.3814168	0.5043901	0.3230146
Antofagasta	0.3993621	0.5214593	0.3305611
Los Ríos	0.3695100	0.5116964	0.3564916
Atacama	0.3185580	0.4204020	0.2500149
Tarapacá	0.3094293	0.3907283	0.2032124
Arica y Parinacota	0.2767496	0.3806655	0.2383202
Magallanes y de La Antártica Chilena	0.2902128	0.4054920	0.2653866
Aysén del Gral. C. Ibáñez del Campo	0.1616789	0.2294529	0.1376292

### IV.3 Prueba de la hipótesis específica 3

A continuación, se detallará los resultados obtenidos con la técnica de series temporales aplicando el modelo ARIMA. En la Figura 4 se muestra el comportamiento de la mortalidad de la región de Santiago desde el año 1997 hasta el 2017.

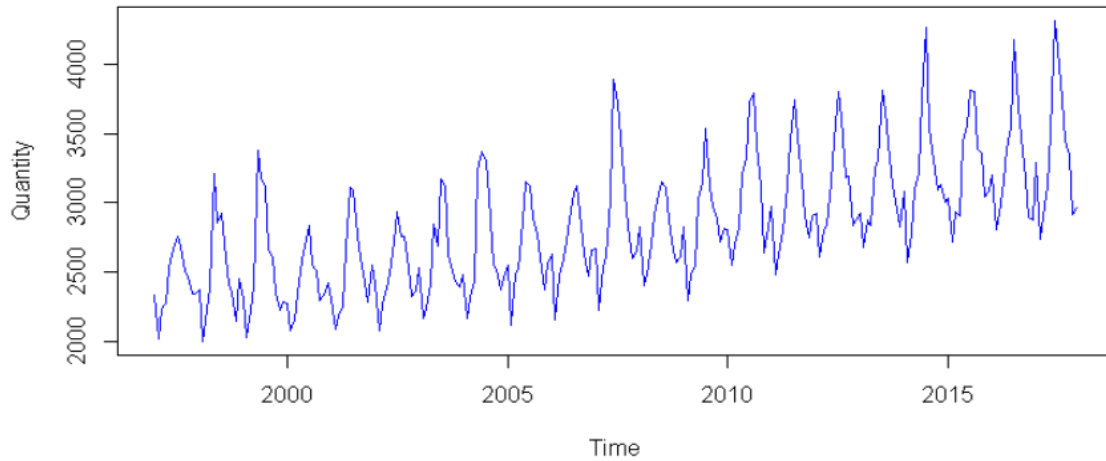


Figura 4. Comportamiento de la mortalidad de Santiago

En la tabla 10 se muestra los resultados de precisión que se obtuvo al aplicar el modelo, donde la métrica más importante es el MAPE, debido a que indica el error porcentual absoluto medio del modelo.

Tabla 10: Resumen del modelo ARIMA

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Santiago	-5.1146	141.4304	98.4152	-0.5057	3.3578	0.6693

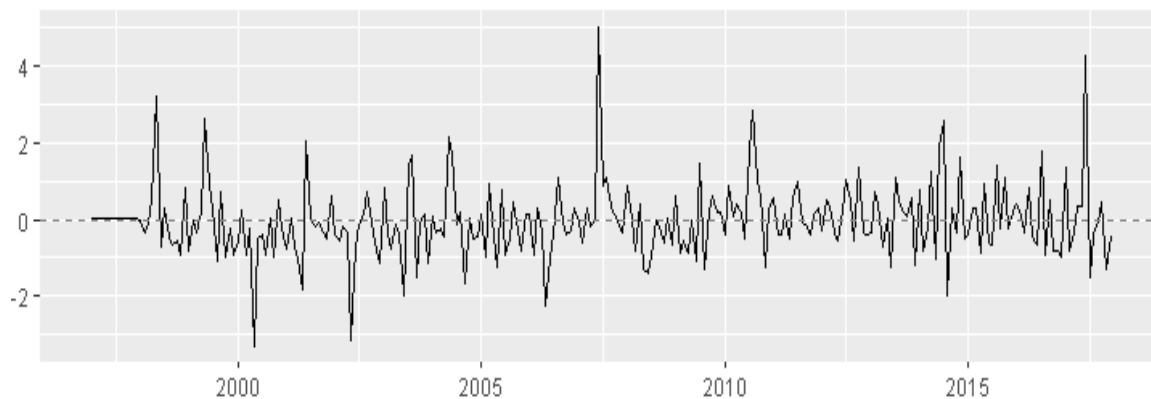
En la tabla 11 se muestra los coeficientes de cada variable obteniendo un modelo estacionario ARIMA (1,0,0) (0,1,1) [12]. Donde los parámetros para reemplazar en la fórmula son:  $p = 1$ ,  $d = 0$  y  $q = 0$ . Por otro lado, para la parte estacionaria:  $P = 0$ ,  $D = 1$  y  $Q = 1$

Tabla 11: Coeficientes del modelo ARIMA

Variable	Estimate	Std, Error	Z value	Pr(> z )
AR1	0.433791	0.058258	7.4461	9.617e-14
SMA1	-0.812078	0.044615	-18.2019	2.2e-16
DRIFT	3.966800	0.338996	11.7016	2.2e-16

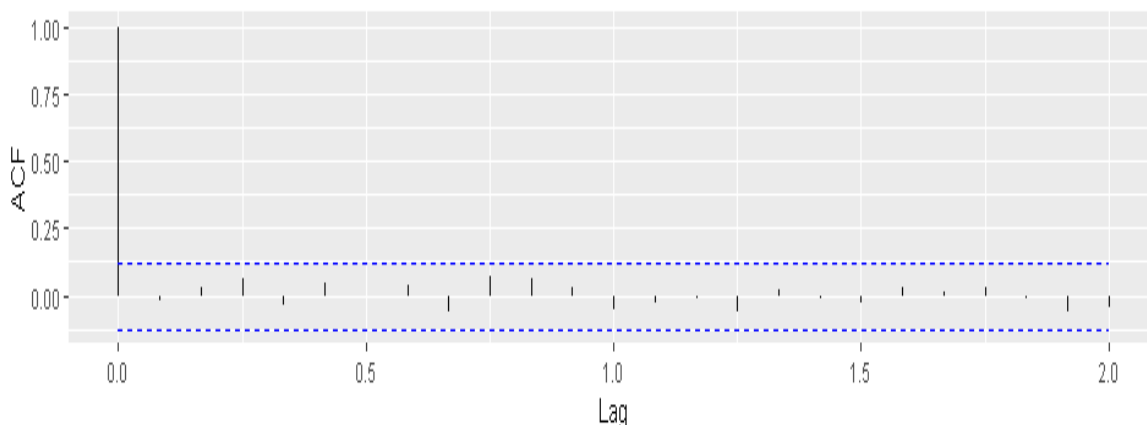
## Diagnóstico del modelo ARIMA

A continuación se detalla las pruebas de autocorrelación de residuos, la prueba de ruido blanco para corroborar que el modelo es bueno y tiene un buen ajuste. Se muestra los valores predecidos por el modelo vs. los valores reales y por último el pronóstico del modelo para el año siguiente. En la figura 5 se muestra los residuos estandarizados del modelo, el cual indica que el error tiene una media igual 0.



*Figura 5.* Residuos estandarizados del modelo ARIMA de Santiago

En la figura 6 se muestra la función de autocorrelación de los residuos donde indica que no existen correlaciones importantes y puede que exista ruido blanco. Posteriormente se podrá corroborar a través de la prueba de Ljung-Box.



*Figura 6.* ACF de los residuos del modelo ARIMA de Santiago



En la figura 7 se muestra la prueba de ruido blanco usando el estadístico de Ljung-Box. Donde,  $H_0$ : Ruido blanco  $> 0.05$  y  $H_1$ : No hay ruido blanco  $< 0.05$ . En el modelo los valores p son mayores a 0.05 por lo cual se acepta la  $H_0$ , el cual indica que existe ruido blanco y el modelo se ajusta bien.

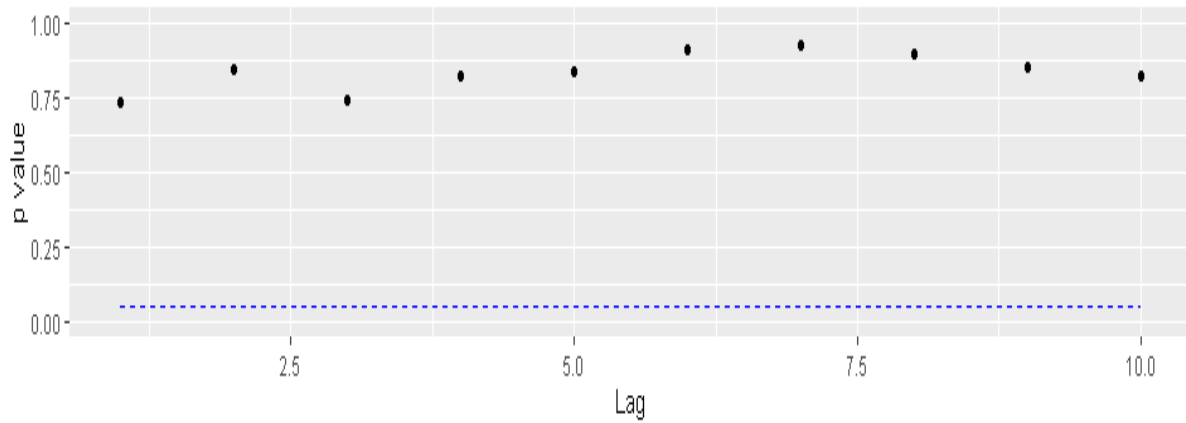


Figura 7. Prueba del ruido blanco del modelo ARIMA de Santiago

En la figura 8 se muestra los valores predecidos con el modelo ARIMA vs. los valores reales, logrando un excelente ajuste a los datos reales.

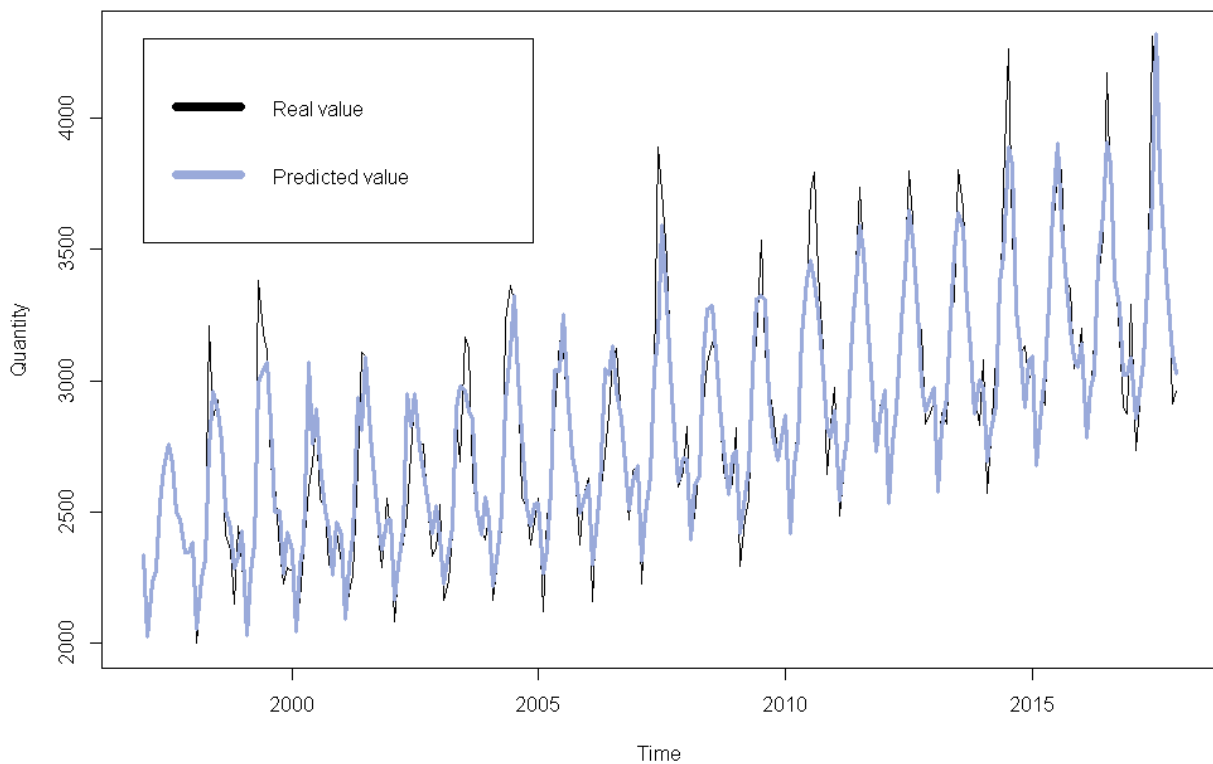


Figura 8. Valores predecidos del modelo ARIMA vs. valores esperados

En la Figura 9 se indica los valores predecidos por el modelo ARIMA para el año posterior. La línea azul es la media de los valores predecidos y las líneas sombreadas de color gris son los intervalos de dicha predicción.

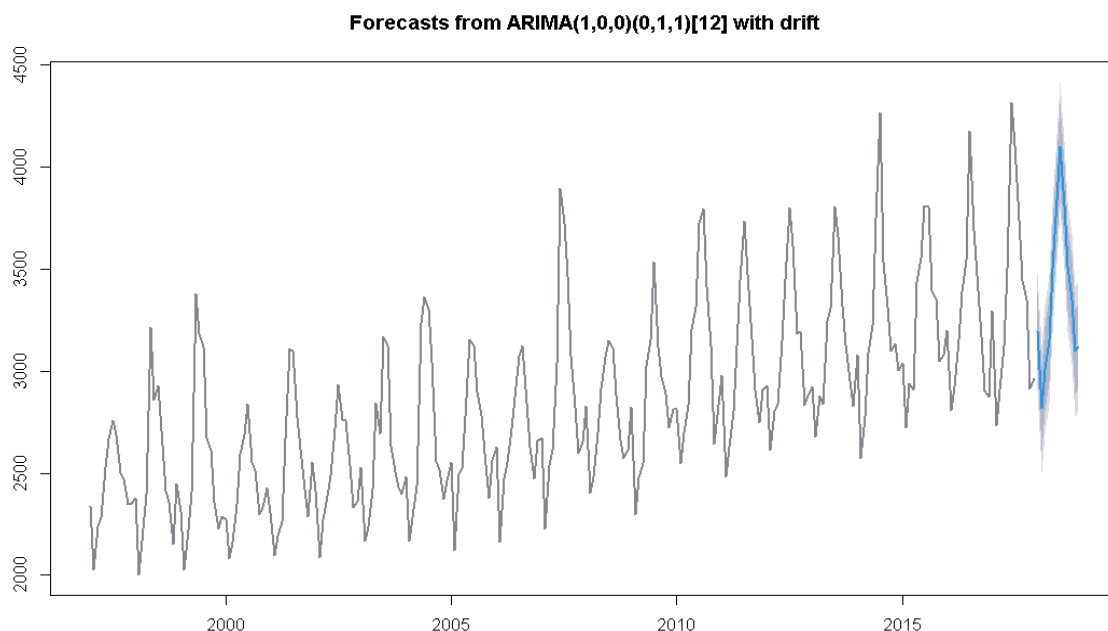


Figura 9. Pronóstico del modelo ARIMA

### Resumen de los modelos ARIMA

El modelo ARIMA fue aplicado a cada una de las 16 regiones de Chile. En la tabla 12 se muestra los resultados obtenidos al aplicar el modelo estacionario ARIMA (1,0,0) (0,1,1) [12] a cada región de Chile ordenado de menor a mayor según el MAPE.

Tabla 12: Resumen de los modelos ARIMA por región

REGIÓN	ME	RMSE	MAE	MPE	MAPE
Metropolitana de Santiago	-5.1146	141.4304	98.4152	-0.5057	3.3578
Valparaíso	-1.8786	47.48408	35.6638	-0.5754	3.9762
Biobío	-1.1608	41.5533	30.2635	-0.5345	4.3288
Maule	8.8537	40.2790	28.8810	1.2864	5.7546
La Araucanía	-1.5589	30.2382	22.1755	-0.7386	4.5570
Libertador General Bernardo O'Higgins	0.2544	26.9128	20.752	-0.4278	5.2084
Los Lagos	9.4401	29.9794	22.8530	2.0986	5.8527

REGIÓN	ME	RMSE	MAE	MPE	MAPE
Coquimbo	-1.1392	21.87694	16.5561	-1.1040	5.7282
Ñuble	-1.2505	19.2729	14.5753	-1.0451	5.7922
Antofagasta	-0.1138	23.0759	18.6217	-0.8233	8.6923
Los Ríos	2.1141	17.8165	13.8879	0.2667	7.0812
Atacama	-1.4040	12.7431	9.7610	-0.0843	9.2153
Tarapacá	1.7380	11.8725	9.2250	0.4675	9.5728
Arica y Parinacota	1.9221	11.5922	9.3648	0.4937	10.930
Magallanes y de La Antártica chilena	1.3333	10.7914	8.4811	-0.1111	10.588
Aysén del Gral. C. Ibáñez del Campo	0.4768	6.9428	5.4345	-1.6048	14.167

## IV.5 Resumen

Tabla 13: Resumen de los resultados de las comprobaciones de las hipótesis

Cod	Hipótesis	Resultado (Aceptada o Rechazada)
<b>HE1</b>	La precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple fue de al menos 80%.	<b>Aceptada</b>
<b>HE2</b>	La precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal fue al menos del 60%.	<b>Aceptada</b>
<b>HE3</b>	El error porcentual absoluto medio del pronóstico de mortalidad con la técnica de series de tiempo fue inferior al 5%.	<b>Aceptada</b>
<b>HG</b>	El uso de técnicas de minería de datos generó mejores resultados en términos de incremento de precisión y reducción del error en el pronóstico de mortalidad.	<b>Aceptada</b>

Conforme se muestra en la tabla 13, en base a los resultados obtenidos se demostró que las hipótesis alternas planteadas inicialmente fueron aceptadas, logrando cumplir el objetivo general y los objetivos específicos. Se generó un incremento de precisión de 82.2% y una reducción del error porcentual absoluto medio de 3.35%.

## **V. DISCUSIÓN**

Las técnicas de minería de datos que fueron aplicadas para el aumento de la precisión y reducción del error porcentual absoluto sobre las defunciones en las regiones de Chile tuvieron resultados positivos. Tanto la regresión lineal múltiple como las series temporales (Arima) tuvieron resultados mejorados con respecto a otras investigaciones, en cambio en la regresión logística ordinal se tuvo un resultado menor a la precisión a otras investigaciones. Conforme a los resultados por regiones, los mayores porcentajes en la técnica de regresión lineal múltiple los tuvieron Valparaíso y Santiago, ya que obtuvieron una precisión de 82.2% y 81.3% respectivamente. En cuanto la técnica de series temporales, las regiones que obtuvieron reducciones del error porcentual fueron: Santiago, Valparaíso y Biobío, ya que con la técnica se obtuvo errores porcentuales de 3.35%, 3.97% y 4.32%, respectivamente.

En base a los resultados obtenidos en la presente investigación se realizaron comparaciones con otras investigaciones, para obtener nuevas ideas sobre el pronóstico de la mortalidad. La técnica de regresión lineal múltiple fue aplicada a una fuente de datos que contiene la tasa de mortalidad de Chile, ya que su propósito fue pronosticar la cantidad de muertes. En esta investigación se establecieron 20 variables independientes, donde las que obtuvieron un mayor impacto fueron las enfermedades del sistema respiratorio (J00-J99), neoplasias (C00-D48) y enfermedades del sistema circulatorio (I00-I99). Las precisiones fueron 23.4 %, 63.6% y 77.8% respectivamente. Por lo contrario, las variables que obtuvieron un menor impacto fueron las enfermedades de la piel (L00-L99), sistema osteomuscular (M00-M99), embarazo, parto y puerperio (O00-O99) con precisiones de 0.3%, 0.3% y 0.2%, respectivamente.

Estas variables independientes se aplicaron respecto a la variable independiente (cantidad de muertes). La mayoría de las variables están relacionadas con las enfermedades de la CIE-10. Esto es importante debido a que esto influyó mucho a la variable causa de muerte. Por otro lado, la región que obtuvo una mejor precisión fue Valparaíso con 82.2%, lo que indica que el modelo es muy eficiente respecto a la región. Por lo contrario, los resultados de este estudio fueron diferentes a los resultados de los estudios de Lin et al. (2015) y Dallolio et al. (2012), quienes encontraron una precisión en los resultados de

90% y 76%, respectivamente. Los resultados del presente estudio fueron diferentes a los obtenidos en la investigación de Lin et al. (2015), en la que se aplicó la técnica de regresión lineal múltiple para modelar y pronosticar tasas de mortalidad en tres diferentes países, donde el modelo contiene solo las variables sexo y edad, en los países de Estados Unidos, Reino Unido y Japón. El resultado fue una precisión mayor al 80%. Se concluyó que la investigación de Lin et al. (2015) solo se basó en la variable edad y sexo, más no incluyeron las enfermedades internacionales.

Con respecto a la investigación de Dallolio et al. (2012), los resultados del presente estudio fueron diferentes porque se basó en la evaluación de la economía del país de Italia y como influyó en la mortalidad infantil. Las variables tomadas en la investigación fueron el ingreso neto, ingreso promedio familiar, porcentaje de mujeres con al menos de 8 años de educación y el porcentaje de desempleados de 15 a 64 años, ya que esta investigación se basó en factores o variables económicas y sociales. Se concluyó que los resultados fueron diferentes porque la investigación de Dallolio et al. (2012) tuvo diferentes objetivos y variables a los utilizados en la presente investigación.

Los resultados del estudio con la técnica de regresión logística ordinal que fue aplicada a una fuente de datos de la mortalidad en Chile se basaron en las probabilidades de cada variable, en síntesis que variable independiente tiene más o menos participación en la variable dependiente. Las variables con mayores probabilidades de muerte fueron las presencias de enfermedades del sistema respiratorio (J00-J99), neoplasias (C00\_D48) y enfermedades del sistema circulatorio (I00-I99), en las que los resultados fueron 180.91, 51.0, 95 y 2.39 veces mayores a los que no pertenecen a estas enfermedades, dejando a las otras variables constantes. En este caso estas variables son las más probables en aumentar la tasa de mortalidad.

Los resultados de esta investigación fueron diferentes a los resultados de los estudios de Hsieh et al.(2018), Adedokun et al.(2020), Singh et al.(2020) e Iriond et al.(2020), quienes encontraron la precisión de los resultados 72.7%, 32% y 40% respectivamente. Los resultados del presente estudio fueron

diferentes a los estudios descritos porque la investigación de Hsieh et al.(2018) se enfoca en desarrollar un modelo de predicción para el cáncer de páncreas y diabetes, por lo tanto los datos de esta investigación solo le dan la prioridad a la información para factores de riesgo de cáncer más no el pronóstico de la mortalidad.

Los resultados del presente estudio fueron diferentes a los resultados de la investigación de Adedokun et al.(2020), quienes se basaron en la mortalidad infantil en Nigeria, donde las principales enfermedades que causaron las defunciones de las personas fueron: neumonía, diarrea y malaria. Adedokun et al.(2020) aplicaron la técnica de regresión ordinal, no obstante las probabilidades se basaron en las variables edad y ciudad. El objetivo también fue diferente, ya que se basaron en la causante de la muerte de los niños y en ciertas enfermedades más constantes en su país.

Los resultados del presente estudio fueron diferentes a los resultados de la investigación de Singh et al.(2020), quienes se basaron en la mortalidad del cáncer oral en India, debido a que este tipo de enfermedad es el más propenso en las personas y la tasa de supervivencia es de 5 años; por lo tanto, la investigación tuvo por objetivo determinar únicamente las muertes sobre un tipo de enfermedad. Los resultados del presente estudio fueron diferentes a los resultados de Iriond et al.(2020), quienes estudiaron la mortalidad infantil. Esta investigación se dividió en 3 modelos de predicción que fueron prenatal, primeras 24 horas de vida y durante su ingreso, donde las precisiones de los modelos fueron de 32% y 40% respectivamente.

Las series de tiempo se crearon utilizando los modelos Arima para pronosticar la cantidad de muertes. En consecuencia, se tuvieron que hacer varios modelos para cada una de las regiones de Chile, debido a que se obtuvo los resultados de cantidad de muertes por región, año y mes. El mejor resultado se dio en la región de Metropolitana Santiago, donde el modelo arrojó un valor respecto al MAPE (error de porcentaje absoluto medio) de 3.357 %. Este modelo ve el error porcentual más no la precisión, pero esto indica que si es menor el porcentaje del error del pronóstico, entonces es mejor la precisión del modelo.

También se utilizó la prueba de Box-Jung para nuestro modelo Arima  $(1,0,0)(0,1,1)_{12}$ , que indica si el modelo Arima se ajusta a los datos del estudio. El ajuste fue 0.7398, que significa la existencia de ruido blanco, modelo media igual a cero, varianza constante y no hay mucha correlación. Estos resultados fueron diferentes a los resultados de los estudios de Lin et al. (2015) y Earnest et al. (2019) y Nayak et al. (2019), los que fueron 11.91 % , 5.2 % y 5% respecto al error porcentual.

Los resultados del presente estudio fueron diferentes a los resultados de la investigación de Lin et al. (2015), ya que dicha investigación tuvo por objetivo predecir la mortalidad por lesiones, es decir que solo se basó en una enfermedad más no en todas las enfermedades de la CIE-10. También crearon tres modelos Arima, ya que tuvieron que medir el resultado y pronosticar la mortalidad por lesiones. Los modelos fueron Arima  $(0,1,1)$ , Arima  $(4,1,0)$  y Arima  $(1,1,(2))$ , en los que los porcentajes de error porcentual respecto a los modelos fueron 11.91%, 11.96% y 11.90%, respectivamente. Los modelos también se ajustaron a la información, ya que existe ruido blanco ( $p > 0.05$ ). El mejor modelo de esta investigación fue Arima  $(0,1,1)$  porque no tenía muchos parámetros.

Los resultados del presente estudio fueron diferentes a los resultados de la investigación de Earnest et al. (2019), ya que se basó en pronosticar el cáncer de próstata en Australia, donde esta enfermedad es una de las más comunes. El mejor resultado para su información fue el modelo Arima  $(1,1,0)$ , ya que su objetivo fue prevenir el cáncer más no la mortalidad. Se utilizó un proceso similar para el modelo de incidencia estandarizada con las variables edad y tasa de mortalidad para el cáncer próstata en Australia durante el mismo periodo de tiempo.

Los resultados del presente estudio fueron diferentes a los resultados del estudio de Nayak et al. (2019), ya que en dicha investigación se estudió la enfermedad epidémica llamada dengue en India. El modelo Arima  $(1,0,0)(1,0,0)_{12}$  fue estacional para pronosticar la carga de la enfermedad, pronosticar tendencias futuras y brotes inminentes. La verificación del ajuste del modelo, en cuanto a la información se hizo con la prueba Ljung-Box donde el resultado fue



de 11.271, que quiere decir que el modelo es adecuado para aplicarlo en esa enfermedad. Esta investigación tuvo como semejanza el modelo Arima  $(1,0,0)(1,0,0)_{12}$ , para pronosticar la tasa de mortalidad, sin embargo el objetivo es distinto, ya que la investigación pronostica la carga de una sola enfermedad, más no el estudio de la mortalidad por varias enfermedades.

## **VI. CONCLUSIONES**

Las conclusiones de la investigación fueron las siguientes:

1. La técnica de regresión lineal múltiple para el pronóstico de la mortalidad es muy bueno, ya que se obtuvieron pronósticos superiores al 70% de precisión en las regiones de Valparaíso, Santiago, Biobío, Maule, O'Higgins y Araucanía, lo que indica que tiene un buen poder explicativo en cuanto la mortalidad.
2. En base a los resultados presentados previamente con la técnica de regresión lineal múltiple, se concluye que la causa de muerte con mayor impacto en la mortalidad en la región de Valparaíso son las enfermedades del sistema circulatorio, debido a que se obtuvo un impacto del 77.8% y la enfermedad que tuvo el menor impacto fueron las relacionadas con el embarazo, el parto y el puerperio con un 0.2%, manteniendo constantes a las otras condiciones.
3. La aplicación de la técnica de regresión lineal múltiple considerando las regiones con un  $r$  cuadrado ajustado superior al 70%, que cuando exista presencia en hospitales o clínicas en las regiones de Valparaíso, Santiago, Biobío, Maule, O'Higgins y Araucanía, las muertes disminuyeron en un 6.9%, 10.8%, 4.5%, 6.9%, 8.2% y 14.4% respectivamente, siendo la región de la Araucanía que la presencia de hospitales o clínicas tiene un impacto del 14.4% en la reducción de la mortalidad.
4. Con la ayuda de la aplicación de la técnica de regresión lineal múltiple considerando las regiones con un  $r$  cuadrado ajustado superior al 70%, que cuando exista presencia femenina en las regiones de Valparaíso, Santiago, Biobío, Maule, O'Higgins y Araucanía, las muertes disminuyeron en un 3.3%, 2%, 7.5%, 9.9%, 10.5% y 8% respectivamente,

siendo la región de O'Higgins, que la presencia del género femenino tiene un impacto del 10.5% en la reducción de la mortalidad.

5. De acuerdo a los resultados obtenidos, se obtuvo que las lesiones, intoxicaciones y algunas otras consecuencias de causas externas tienen un impacto entre el 19.4% y el 25.5%, siendo la región de la Araucanía con el mayor impacto del 25.5% de aumento en la mortalidad. Eso significa que se debería prestar más atención a la causa de muerte que afecta más en esa región.
6. Aplicando la regresión ordinal, se puede decir que los hombres tienen más probabilidades de morir, puesto que los resultados que mostraron la variable sexo con respecto a las mujeres tenían una probabilidad de morir 1.01 veces, lo que significa que tienen la probabilidad de morir en 1 %, esto quiere decir que la probabilidad de morir con la variable sexo en relación con los hombres es muy alta.
7. En los resultados de la regresión ordinal, las enfermedades que tenían más probabilidades de causar muertes en la región de Valparaíso eran enfermedades del sistema circulatorio, enfermedades del sistema respiratorio y neoplasias que tenían un resultado de probabilidad 180.91 (0.91%) veces, 2.392885596 (0.39%) veces y 51.09 (0.095%) veces, es decir mayores problemas de muerte que otras enfermedades, de lo contrario con las otras enfermedades que tienen más del 95% de probabilidades de que las personas no mueran por estas causas.
8. En la regresión ordinal aplicada a la región de Valparaíso, se concluyó que las probabilidades de muerte por rango de edad de 0 a 64 tienen una probabilidad mayor de 98% de que la mortalidad no ocurra en las personas.
9. La técnica de series temporales usando el modelo ARIMA es bueno, ya que se obtuvo un MAPE inferior al 5% en las regiones de Santiago, Valparaíso y Biobío con un MAPE de 3.35%, 3.97% y 4.32%

respectivamente y un valor en la prueba de ruido blanco de 0.8549, 0.8549 y 0.6607 respectivamente, que es superior a 0.05, lo que indica que los modelos están bien ajustados.

10. Los modelos ARIMA tradicionales no cumple como la mejor opción para pronosticar la mortalidad, esto se demuestra porque en todas las regiones el mejor modelo incluía la estacionalidad, como por ejemplo en la región de Santiago con un modelo ARIMA (1,0,0) (0,1,1) con 12 períodos, que arrojó el MAPE más bajo de todos.
11. La metodología CRISP-DM obtuvo una puntuación más alta porque es la más actual de todas las metodologías de minería de datos, era flexible en todas las fases como recopilar datos, explorar los datos, limpiar los datos y crear modelos para cada técnica de minería de datos y para la implementación del GIS. En otras palabras, esta metodología ayudó a proyectar la predicción en un sistema geolocalizado.
12. El uso de Azure Machine Learning para la construcción de modelos de minería de datos es bastante bueno, ya que tiene varias técnicas de minería, además de módulos para la programación en lenguaje R o Python, esta herramienta se utilizó debido a su fácil y rápido uso, ya que es un servicio en la nube y no necesita tener una buena computadora para procesar grandes cantidades de datos.
13. El uso de ArcGIS Online es muy bueno, debido a que sirvió como un proyector de mapas de los resultados obtenidos por las técnicas de minería de datos, la suma de estas herramientas ayuda a tener una visión clara de la mortalidad por región, esto ayudará a tomar decisiones y crear estrategias para reducir la mortalidad.
14. El uso de la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo predicen la mortalidad de una forma correcta, debido a que todas las hipótesis propuestas en la investigación fueron aceptadas. Además, que la enfermedad que tiene más presencia en todas las regiones de Chile son enfermedades del sistema circulatorio con un

impacto promedio del 72.85%. Por otro lado, la presencia de hospitales o clínicas llega a reducir hasta un 14.4%, lo que indicó que cuanto mayor es la presencia de hospitales y clínicas, la mortalidad reduce un buen porcentaje.

## **VII. RECOMENDACIONES**

Las recomendaciones para futuras investigaciones son las siguientes:

1. Crear modelos relacionados con los grupos de enfermedades del ojo y sus enfermedades anexas (H00-H59) y enfermedades del oído y del proceso mastoideo (H60-H95), debido a que es muy importante saber cuáles son los factores que puedan influir en la muerte por estos grupos de enfermedades.
2. Aplicar modelos de pronóstico para predecir la mortalidad por grupos de enfermedades. No existe un modelo mejor que prevalezca sobre los demás, lo que sí existe es un modelo más adecuado para cada caso de muerte y con esto podrían tomar decisiones para evitar futuras muertes.
3. Aplicar diferentes técnicas de series de tiempo para poder pronosticar la mortalidad, ya que en esta investigación solo se propuso predecir el número de muertes por región, dado que se dio un buen resultado, no se eligió hacer una comparación con otras técnicas de series de tiempo.
4. Realizar un estudio donde sea aplicado por grupos de enfermedades, es decir, el número de muertes se puede visualizar mediante una técnica estadística, pero con la variación de que sea analizado por año y mes, con el fin de saber cómo atacará a estos grupos de enfermedades en los próximos años.
5. Desarrollar una investigación científica considerando una mayor especificidad en la clasificación CIE-10, ya que la presente investigación se centró en el análisis por capítulos, pero cada capítulo tiene grandes grupos de enfermedades. Esto ayudaría a tener un análisis detallado o un modelo que identifique el patrón de mortalidad para cada grupo, categoría o enfermedad de la CIE-10.
6. Llevar a cabo investigaciones científicas en un área geográfica mayor a la utilizada en esta investigación, ya que los modelos se realizaron por región de Chile, pero sería muy útil crear modelos para cada unidad básica de la



estructura político-administrativa de un país, con la finalidad de tener un mejor patrón de comportamiento de mortalidad, ya que cada ubicación territorial tiene un comportamiento diferente

7. Utilizar otras técnicas de minería de datos que podrían conducir a una mejor interpretación de los datos procesados en futuras investigaciones, como por ejemplo el uso de Machine Learning o Deep Learning, porque tiene algoritmos más desarrollados para encontrar patrones de comportamiento.
8. Ampliar la investigación científica con una muestra más grande de documentos para evaluar. En este estudio solo se utilizaron revistas y libros científicos en idioma ruso e inglés y con un período máximo de 5 años, esto podría extenderse a más idiomas o períodos de tiempo para encontrar datos relevantes para el pronóstico de la mortalidad.
9. Como una línea de trabajo futura, para tratar de medir cuantitativamente el manejo y evaluar los logros y metas a casi el 100%, se propone analizar modelos predictivos sobre mortalidad, nacimiento y calidad de vida de los pacientes.

## **REFERENCIAS**

ABED, Amira y NASR, Mona. Diabetes Disease Detection through Data Mining Techniques. International Journal of Advanced Networking and Applications [online]. 2019, vol, 11. No, 1. pp, 4142-4149. ISSN 09750290.

ADEDOKUN, Sulaimon. Correlates of childhood morbidity in Nigeria: Evidence from ordinal analysis of cross-sectional data. PLoS One [online]. 2020, vol,15. No, 5.

AHMED. Analysis of Data Mining Tools for Disease Prediction. Journal of Pharmaceutical Sciences and Research [online]. 2017, vol. 9, No.10, pp. 1886-1888.

ALBY,S y SHIVAKUMAR, B.L. A Novel Approach for Prediction of Type 2 Diabetes. International Journal of Advanced Research in Computer Science [online]. 2016, vol. 7, no. 4.

ALPA, R, JAYESH, J. Data modeling techniques for data warehouse. International Journal of Multidisciplinary Research, 2012, vol. 2, no 2, p. 240-246.

ANGADI, Mahantesh y KULKARNI. Amogh. Time Series Data Analysis for Stock Market Prediction Using Data Mining Techniques with R. International Journal of Advanced Research in Computer Science [online]. 2015, vol. 6, No. 6.

ARIAS, Fidas. El proyecto de investigación. Introducción a la metodología científica, 6.a ed. Venezuela: Editorial Episteme, 2012, pp.143.

ARRIETA, Alí, FLÓREZ-TANÚS, Alvaro y ALVIS-GUZMÁN, Nelson. Impacto de la mortalidad evitable en los patrones de mortalidad de la Región Caribe, 1999-2014. Revista De Economía Del Caribe [online]. 2018, No. 21, pp. 25-54. ISSN 20112106.

AWAD, Aya, et al. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *International Journal of Medical Informatics* [en línea]. 2017, 108, 185–195. ISSN 1386-5056.

BASSON, Lisa, KILBOURN, Peter y WALTERS, Jackie. Forecast accuracy in demand planning: A fast-moving consumer goods case study. *Journal of Transport and Supply Chain Management* [online]. 2019, vol. 13. ISSN 23108789.

BERGERON-BOUCHER, Marie-Pier. Coherent modeling and forecasting of mortality patterns for subpopulations using multiway analysis of compositions: An application to Canadian provinces and territories. *North American Actuarial Journal*. 2018, vol. 22, No.1, p. 92-118.

BIBLIOTECA DEL CONGRESO NACIONAL DE CHILE. Biblioteca del Congreso Nacional | Ley Chile. [www.bcn.cl/leychile](http://www.bcn.cl/leychile) [en línea]. 2020. Disponible en: <https://www.bcn.cl/leychile/navegar?idNorma=141599>.

BOWERS, William y JARRELL, Stephen. Accuracy Analysis of Consensus Forecasts. *The Journal of Business Forecasting*. 1992, vol. 11, No.2, p. 9.

CHEN, Feng, et al. Data Mining for the Internet of Things: Literature Review and Challenges. *International Journal of Distributed Sensor Networks* [online]. 2015. ISSN 15501329.

CHENG, Xiaoqing, et al. A multiple linear regression model for structure of n-linked oligosaccharides [online]. Stevenage: The Institution of Engineering & Technology, 2015. Copyright - Copyright The Institution of Engineering & Technology Aug 21, 2015; Última actualización - 2019-03-01. ISBN 978-1-78561-085-1.

CHINCHILLA, Ricardo. Mercado de datos: conceptos y metodologías de desarrollo. *Tecnología en Marcha*, 2011, vol. 24, no 3, p. 55-66.

CORNEJO, David. Estudio del comportamiento de la mortalidad en Chile basado en la minería de datos y análisis predictivos. 2017, Tesis Doctoral. PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO.

DADERMAN, Antonia y ROSANDE, Sara. Evaluating frameworks for implementing machine learning in signal processing: A comparative study of CRISP-DM. *semma and kdd*, 2018.

DALLOLIO, Laura. Socio-economic factors associated with infant mortality in Italy: an ecological study. *International Journal for Equity in Health* [online]. 2012, vol. 11, pp. 45.

EARNEST, Arul. Forecasting annual incidence and mortality rate for prostate cancer in Australia until 2022 using autoregressive integrated moving average (ARIMA) models. *BMJ Open* [online]. 2019, vol. 9, No. 8.

ENRÍQUEZ, Elvis. Sistema de información web y su mejora en la gestión académica del colegio privado Hans Kelsen del distrito de Florencia de Moratrujillo. En: <https://dspace.unitru.edu.pe/> [base de datos en línea]. Para optar el Título profesional de Ingeniero de Sistemas, Universidad Nacional de Trujillo, 2016.

ER, Anita y ER, Jasjeet. A Survey on Data Mining and Its Current Research Directions. *International Journal of Advanced Research in Computer Science* [online]. 2017, vol. 8, No. 4.

FASHOTO, Stephen. MODELLING OF THE FEED FORWARD NEURAL NETWORK WITH ITS APPLICATION IN MEDICAL DIAGNOSIS. *International Journal of Advances in Engineering & Technology* [online]. 2015, vol. 8, No. 4, pp. 507-520.

FAYYAD, Usama, PIATETSKY-SHAPIRO, Gregory y SMYTH, Padhraic. From data mining to knowledge discovery in databases. *AI magazine*. 1996, vol. 17, No. 3, p. 37.

FEKI, Jamel, HACHAICHI, Yasser. Constellation discovery from OLTP parallel-relations. En Proceedings of the 8th International Arab Conference on Information Technology ACIT. 2007.

FOOTE, Kennet y LYNCH, Margaret. Geographic Information Systems as an Integrating Technology. GISWeb [enlínea]. Disponible en: <http://gisweb.massey.ac.nz/topic/webreferencesites/whatisgis/texaswhatisgis/texas/intro.htm>.

GAETA, M, et al. European cardiovascular mortality over the last three decades: evaluation of time trends. forecasts for 2016. 2017.

GARANI, G. and HELMER, S. Integrating Star and Snowflake Schemas in Data Warehouses. International Journal of Data Warehousing and Mining, 2012, vol. 8, no. 4. pp. 22 ProQuest Central. ISSN 15483924.

GIM, Jangwon, LEE, Sukhoon y JOO, Wonkyun. A Study of Prescriptive Analysis Framework for Human Care Services Based On CKAN Cloud. Journal of Sensors [online]. 2018, vol. 2018, pp. 10. ISSN 1687725X.

GIRALDO, Diana, et.al. A model for mortality analysis in Colombia. 2000-2012. Revista De Salud Publica [online]. 2017, vol. 19, No. 2, pp. 241-249. ISSN 01240064.

GONZÁLEZ , María José. Metodología de la investigación social: Técnicas de recolección de datos. Alfabara. Madrid (España). 1997. ISBN 84-8018-108-7.

GONZÁLEZ, María. Análisis de series temporales: Modelos ARIMA. Vasco: Universidad del País Vasco. 2009. ISBN 978-84-692-3814-1.

HAND, David J.; HENLEY, William E. Statistical classification methods in consumer credit scoring: a review. Journal of the Royal Statistical Society: Series A (Statistics in Society), 1997, vol. 160, no 3, p. 523-541.

HERNÁNDEZ, Akram, DÍAZ, Deysi, ESPINOZA, Daniel y VILCARROMERO, Stalin. Análisis espacial de la mortalidad distrital por enfermedades cardiovasculares en las provincias de Lima y Callao. Revista Peruana de Medicina Experimental y Salud Pública [en línea]. 2016, 33(1), 185. ISSN 1726-4642.

HERNÁNDEZ, Roberto, FERNÁNDEZ, Carlos y BAPTISTA, Maria. Metodología de la Investigación, 6.a ed. México DF: Mc Graw Hill. 2014, pp. 599.

HILL, Robin. What an Algorithm Is. Philosophy & Technology [online]. 2016, vol. 29, No. 1, pp. 35-59. ISSN 22105433.

HOPPING, Clare, HELLARD, Bobby y MILLMAN, Rene. What is data and big data mining? an easy guide [online]. London: Dennis Publishing Ltd. 2018. Copyright - Copyright Dennis Publishing Ltd. Oct 16. 2018; Última actualización - 2018-10-17.

HSIEH, Meng, et.al. Development of a prediction model for pancreatic cancer in patients with type 2 diabetes using logistic regression and artificial neural network models. Cancer Management and Research [online]. 2018, vol. 10, pp. 6317-6324.

HUERTA, Luis, et al. On The Data Mining Process for Classification of Fetal Death Causes. International Journal of Computer Science and Information Security, 2015, vol. 13, no 8, p. 5.

IBM. Esquemas dimensionales. s.f.  
<https://www.ibm.com/docs/es/ida/9.1.2?topic=design-dimensional-schemas>

IONITA, Irina. DATA MINING FOR PREDICTING THE MILITARY CAREER CHOICE. Land Forces Academy Review [online]. 2015, vol. 20, No. 3, pp. 297-306.

IRIONDO, Martí, et.al. Prediction of mortality in very low birth weight neonates in Spain. PLoS One [online]. 2020, vol. 15, No. 7. ISBN: 978-1-4562-2396-0

ISLAM, Md Saiful, et al. A systematic review on healthcare analytics: application and theoretical perspective of data mining. En Healthcare. Multidisciplinary Digital Publishing Institute. 2018. pp. 54.

KHAN, Shafat; AAQIB, Syed Mutahar. Empirical Evaluation of ArcGIS with Contemporary Open Source Solutions-A Study. International Journal of Advance Research in Science and Engineering, 2017, vol. 6, no 1, p. 724-736.

KÖPPELOVÁ, Jana y JINDROVÁ, Andrea. Application of Exponential Smoothing Models and Arima Models in Time Series Analysis from Telco Area. AGRIS on-Line Papers in Economics and Informatics [online]. 2019, vol. 11, No. 3, pp. 73-84.

KÖPPELOVÁ, Jana y SVATOŠOVÁ, Libuše. Development trends of foodstuffs consumption in CR considering healthy eating. AGRIS on-Line Papers in Economics and Informatics [online]. 2019, vol. 11, No. 3, pp. 45-53.

KUMAR, Mahesh. Soil Erosion Estimation In ArcGIS.2017.

KUMAR, Surender y KAUR, Kanwaldip. Review of Data mining (Knowledge discovery) in the Future. International Journal of Advanced Research in Computer Science [online]. 2016, vol. 7, No. 6.

KUMAR, Vijay y GARG, B.S. Improving Cause of Death Information for Infant Deaths by Application of Verbal Autopsy Tool in India. Indian Journal of Public Health Research & Development [online]. 2013, vol. 4, No. 2, pp. 272-277. ISSN 09760245.

LI, Liling; SHRESTHA, Sharad; HU, Gongzhu. Analysis of road traffic fatal accidents using data mining techniques. En 2017 IEEE 15th International



Conference on Software Engineering Research, Management and Applications (SERA). IEEE, 2017. p. 363-370.

LICO, Ledion. Data Mining Techniques in Database Systems. Journal EJSDR- European Journal of Sustainable Development and Research [online]. 2017, vol. 2. No. 1, pp. 43-50.

LIN, Tzuling y TSAI, Cary. A Simple Linear Regression Approach to Modeling and Forecasting Mortality Rates. Journal of Forecasting [online]. 2015, vol. 34, No. 7, pp. 543. ISSN 02776693.

LIN, Yilan, et.al. Application of an autoregressive integrated moving average model for predicting injury mortality in Xiamen. China. BMJ Open [online]. 2015, vol. 5, No. 12.

LONGLEY, Paul, et.al. Geographic information systems and science. John Wiley & Sons. 2005.

LÓPEZ-ROLDÁN, Pedro y FACHELLI, Sandra. Metodología de la investigación social cuantitativa. 2015.

MOHIUDDIN, Ahmed. An Investigation of Performance Analysis of Anomaly Detection Techniques for Big Data in SCADA Systems. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems [online]. 2015, vol. 2, No. 3.

MONTGOMERY, Douglas, JENNINGS, Cheryl y KULAHCI, Murat. Introduction to time series analysis and forecasting. John Wiley & Sons. 2015.

NAYAK, M. y NARAYAN, Ka. Forecasting Dengue Fever Incidence Using ARIMA Analysis. International Journal of Collaborative Research on Internal Medicine & Public Health [online]. 2019, vol. 11, No. 3, pp. 924-932.

NETO, Cristiana, BRITO, Maria, LOPES, Vítor, PEIXOTO, Hugo, ABELHA, António, MACHADO, José. Application of Data Mining for the Prediction of Mortality and Occurrence of Complications for Gastric Cancer Patients. Entropy [online]. 2019, vol. 21, No. 12, s. 1163.

ÑAUPAS , Humberto et al. Metodología de la investigación cuantitativa-cualitativa y redacción de la tesis. Bogotá: Ediciones de la U, 2018. ISBN 978-958-762-877-7.

OLIVEIRA, Má. On Business Analytics: Dynamic Network Analysis for Descriptive Analytics and Multicriteria Decision Analysis for Prescriptive Analytics. Ann Arbor: Universidade do Porto (Portugal), 2016, Order No. 10593636. ISBN 9781083590077.

Organización Mundial de la Salud (OMS). Clasificación estadística internacional de enfermedades y problemas de salud relacionados, décima revisión. Disponible en: <https://icd.who.int/browse10/2019/en>.

Organización Mundial de la Salud (OMS). Clasificación internacional de enfermedades y problemas de salud relacionados (CIE). 2019. Disponible en: <http://www.who.int/classifications/icd/en/>.

Organizacion Mundial de la Salud (OMS). Las diez principales causas de muerte. Disponible en: Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death>.

Organizacion Mundial de la Salud (OMS). Mortalidad. Disponible en: <https://www.who.int/topics/mortality/es>.

PÉREZ, Edgardo y MEDRANO, Leonardo. Análisis factorial exploratorio: bases conceptuales y metodológicas. Revista Argentina de Ciencias del Comportamiento (RACC). 2010, vol. 2, No 1, p. 58-66.

POUDEL, Krishna , et.al. Prediction of the cancer incidence in Nepal. Asian Pacific journal of cancer prevention: APJCP. 2017, vol. 18, No.1, p. 165.

POWERS, David. Evaluation: From precision. recall and F-measure to ROC. informedness. markedness and correlation [online]. Ithaca: Cornell University Library. arXiv.org. 2020.

RAMOS, Salvador. Data Warehouse, Data Marts y Modelos Dimensionales. Un pilar fundamental para la toma de decisiones. 2a ed. Alicante: SolidQ Global S.A., 2016. ISBN 978-84-940719-2-8.

SAIMA, Anwar, et.al. Application of data mining techniques for medical data classification: A review [online]. Les Ulis: EDP Sciences. 2018.

SAS Institute. SAS Help Center . 2017. Disponible en: <https://documentation.sas.com/doc/en/emref/14.3/n061bzurmej4j3n1jnj8bbj m1a2.htmmej4j3n1jnj8bbj m1a2.htm>.

SHARMA, Richa, SINGH, Shailendra y KHATRI, Sujata. Medical data mining using different classification and clustering techniques: A critical survey [online]. Piscataway: The Institute of Electrical and Electronics Engineers. Inc. (IEEE). 2016, pp. 687-691.

SINGH, Vishwajeet, SADA, Nand y DEO, S. Ordinal logistic regression model describing factors associated with extent of nodal involvement in oral cancer patients and its prospective validation. BMC Medical Research Methodology [online]. 2020, vol. 20, pp. 1-8.

SYAIDATUS, Ahmad, et.al. A Review on Student Attrition in Higher Education Using Big Data Analytics and Data Mining Techniques. International Journal of Modern Education and Computer Science [online]. 2019, vol. 10, No. 8, s. 1. ISSN 20750161.

ȚARANU, Ionut. Data mining in healthcare: decision making and precision. Database Systems Journal. 2016, vol. 6, No. 4, p. 33-40.

VIDAL, Carolina, HOFFMEISTER, Lorena y BIAGINI, Leandro. Tendencia de la mortalidad por cáncer de cuello uterino en Chile: aplicación de modelos de regresión joinpoint. Revista Panamericana de Salud Pública. 2013, vol. 33, p. 407-413.

Wang, W., Du, X. & Xie, C. Road landslide information management and forecasting system base on GIS. Environ Monit Assessment 156, 391. 2009. Disponible en: <https://doi.org/10.1007/s10661-008-0492-2>.

WIRTH, Rüdiger y HIPPEL, Jochen. CRISP-DM: Towards a standard process model for data mining. En Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. London. UK: Springer-Verlag. 2000.

XIE, Liming. Time Series Analysis and Prediction on Cancer Incidence Rates. Journal of Medical Discovery [online]. 2017, vol. 2, No. 3, pp. 1-10.

YOO, Illhoi, et al. Data Mining in Healthcare and Biomedicine: A Survey of the Literature. Journal of Medical Systems [online]. 2012, vol. 36, No. 4, pp. 2431-48. ISSN 0148-5598.

## Anexo 1: Matriz de operacionalización de variables

Tabla 14: *Matriz de operacionalización de variables*

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador	Instrumento	Escala de medición
Efecto de un sistema de pronóstico geolocalizado que aplica técnicas de minería de datos para pronosticar la mortalidad	Según Enríquez (2016), un sistema informático es un conjunto de recursos tecnológicos, humanos y económicos, interrelacionados entre sí, para satisfacer las necesidades de información de una institución, para una gestión y toma de decisiones adecuadas. En este proyecto de investigación, el sistema informático se basa en:	Uso de grandes bases de datos para encontrar y generar patrones repetitivos que ayudan a obtener información adecuada cuyas métricas se expresan en funcionalidad (nivel de precisión de pronóstico) y eficiencia (reducción de error porcentual absoluto medio).	Precisión de pronóstico (Powers. 2007. p. 2)	Precisión incrementada	Pestaña Recopilación de datos	Escala de Valor
	<p>1. Sistema geolocalizado: Longley, Goodchild, Maguire y Rhind (2005) indicaron que una definición interesante que enfatiza el poder de un SIG como motor de análisis, para examinar datos y revelar sus patrones, relaciones y anomalías, cosas que podrían no ser evidentes para alguien mirando un mapa (p. 16).</p> <p>2. Minería de datos: Lico (2017) indica: La minería de datos ofrece numerosas formas de descubrir patrones ocultos en grandes cantidades de datos. Estos patrones ocultos pueden usarse potencialmente para predecir el comportamiento futuro. Los buenos datos son el primer requisito para una buena exploración de datos. Existen varias técnicas y algoritmos que se pueden usar para realizar la extracción de datos, pero su uso depende de la aplicación. Las técnicas de minería de datos predictivas son apropiadas cuando tiene un valor objetivo específico que desea predecir en sus datos. (p. 49)</p>			Reducción de MAPE	Pestaña Recopilación de datos	Escala de Valor
			Error porcentual absoluto medio (Kilborn y Walters. 2019. p.5)	(Hand y Henley. 1997)	(Bowers y Jarrel. 1992. p. 10)	

## Anexo 2: Matriz de Consistencia

Tabla 15: *Matriz de Consistencia*

PROBLEMAS	OBJETIVOS	HIPÓTESIS	VARIABLE	DIMENSIONES	INDICADORES
General	General	General			
¿Cuál fue la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo?	Determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple, regresión logística ordinal y series de tiempo.	El uso de técnicas de minería de datos generó mejores resultados en términos de incremento de precisión y reducción del error en el pronóstico de mortalidad	-	-	-
Específico	Específico	Específico			Indicadores
¿Cuál es la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple?	Determinar la precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple.	La precisión del pronóstico de mortalidad con la técnica de regresión lineal múltiple fue de al menos 80% (Lin y Tsai, 2015, p. 1; Poudel et al., 2017, p. 167).			Precisión incrementada (Hand y Henley, 1997)
¿Cuál es la mejora en la precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal?	Determinar la precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal.	La precisión del pronóstico de mortalidad con la técnica de regresión logística ordinal fue de al menos 60% (Hsieh et al., 2018, p. 6; Baucells et al., 2020, p. 10).		Precisión de pronóstico (Powers, 2007, p. 2)	Precisión incrementada (Hand y Henley, 1997)
¿Cuál es la mejora en la precisión del pronóstico de mortalidad con la técnica de series de tiempo?	Determinar la precisión del pronóstico de mortalidad con la técnica de series de tiempo.	El error porcentual absoluto medio del pronóstico de mortalidad con la técnica de series de tiempo fue inferior al 5% (Lin et al., 2015, p. 7; Nayak y Narayan 2019, p. 5; Xie, 2017, p.1; Köppelová y Svatošová, 2019, p.47; Köppelová y Jindrová, 2019, p. 78).	Efecto de un sistema de pronóstico geolocalizado que aplica técnicas de minería de datos para pronosticar la mortalidad. (Lico, 2017, p. 49; Brockwell y Davis, 2010, p. 56)	Error porcentual absoluto medio (Basson, Kilbourn y Walters, 2019, p. 5)	Reducción de MAPE (Bowers y Jarrell, 1992, p. 10)

## Anexo 3: MODELOS, TÉCNICAS Y METODOLOGÍAS DE MINERÍA DE DATOS

### MODELOS, TÉCNICAS Y METODOLOGÍAS DE LA MINERÍA DE DATOS

La minería de datos proporciona muchas formas de descubrir patrones ocultos en grandes cantidades de datos. Existen varias técnicas y algoritmos que se pueden usar para realizar la extracción de datos, pero su uso depende de la aplicación. Por lo tanto, Lico (2017) indicó: "Hay dos modelos importantes en la minería de datos: el modelo descriptivo y el modelo predictivo" (p. 43). Del mismo modo, existe un modelo híbrido entre los dos mencionados anteriormente, llamado prescriptivo.

- **Modelo descriptivo:** Abed y Nasr (2019), explicaron que el modelo descriptivo explora las propiedades y características de los datos examinados, no está sujeto a la predicción de nuevas características e incluye varios métodos como agrupación, reglas de asociación, resúmenes y análisis de patrones y secuencia. La minería de datos descriptiva está diseñada básicamente para generar frecuencia y secuencia, tabulación cruzada y correlación interesante. Los modelos de minería de datos descriptivos se pueden definir para encontrar regularidades desconocidas interesantes en los datos, para descubrir patrones ocultos y para encontrar un conjunto interesante de subgrupos en la mayor parte de los datos admitidos. (p. 4143)
- **Modelo predictivo:** Ionita (2015) indicó que la predicción de minería de datos es un enfoque que implica descubrir los patrones más poderosos en grandes bases de datos, patrones que pueden generalizar decisiones futuras correctas. El modelo clásico para los datos de predicción son los casos muestreados. Las posibles medidas llamadas características (atributos) son conocidas (especificadas) y medidas en varios casos (situaciones). El papel del modelo de minería de datos predictivo es aprender los criterios de toma de decisiones para asignar etiquetas a otros casos nuevos no clasificados. (p. 299)

- **Modelo prescriptivo:** Gim, Lee Y Joo (2018) indicaron que el modelo prescriptivo es un conjunto de técnicas matemáticas para determinar objetivos complejos, requisitos y limitaciones para mejorar los resultados comerciales (p. 2). El análisis prescriptivo, como un tipo de técnica de análisis de datos, proporciona predicciones específicas del contexto e información personalizada (Gim et al., 2018, p. 2). Esta técnica se utiliza para respaldar una toma de decisiones con múltiples ideas más efectiva cuando los tomadores de decisiones empresariales, como los CTO y CEO, analizan y predicen situaciones complejas (Gim et al., 2018, p. 2).

Lico (2017) detalló: "Hay 4 técnicas importantes utilizadas en el modelo predictivo: clasificación, regresión, series de tiempo y detección de anomalías" (p. 46). Cada técnica se explicará a continuación:

- **Clasificación:** Ahmed y Kauser (2017) indicaron que la clasificación es una técnica de minería de datos basada en el aprendizaje automático. La clasificación se utiliza para clasificar cada información en un conjunto de datos en uno de un conjunto predefinido de grupos o clases. Use técnicas matemáticas como árboles de decisión, programación lineal, redes neuronales y estadísticas para clasificar los datos en diferentes grupos. Las técnicas modernas de clasificación proporcionan métodos más inteligentes para la predicción efectiva de la enfermedad. (p. 1887)
- **Regresión:** Lico (2017) explicó que la regresión es una técnica de minería de datos (aprendizaje automático) utilizada para ajustar una ecuación a un conjunto de datos. Se usa con mayor frecuencia cuando el atributo de destino tiene un valor numérico. Se puede usar para modelar la relación entre una o más variables independientes y variables dependientes. En la minería de datos, las variables independientes ya son atributos conocidos y las variables de respuesta son lo que queremos predecir. (p. 46)
- **Series de Tiempo:** Angadi y Kulkarni (2015) detallaron que las técnicas de análisis de datos de series temporales utilizan información verificable como premisa para evaluar los resultados futuros. Los datos de series de



tiempo pueden definirse como datos numéricos recopilados en una secuencia particular durante un período de tiempo a intervalos regulares. Los datos de series de tiempo pueden incluir los valores recopilados al final de cada semana, mes, trimestre o año, etc. La intención es determinar si existe algún vínculo entre los datos recopilados hasta el momento y cómo cambian los datos. (p. 104)

- **Detección de anomalías:** Mohiuddin, Anwar, Abdun, Shah y Maher (2015) expresaron que la detección de anomalías es una tarea importante del análisis de datos. El objetivo principal de la detección de anomalías es detectar datos anormales o anormales de un conjunto de datos. Esta es un área interesante de investigación de minería de datos, ya que implica descubrir patrones nuevos y raros a partir de un conjunto de datos. La detección de anomalías se ha estudiado ampliamente en estadística y aprendizaje automático. (p. 3)

Los especialistas del Instituto SAS (2017) explicaron que la minería de datos, como el proceso de muestreo, exploración, modificación, modelado y evaluación (SEMMA) de grandes cantidades de datos para descubrir patrones previamente desconocidos eso puede ser usado como una ventaja comercial. El proceso de minería de datos es aplicable en una variedad de industrias y proporciona metodologías para problemas comerciales tan diversos como detección de fraude, hogar, retención y agotamiento de clientes, marketing de bases de datos, segmentación de mercado, análisis de riesgos, análisis de afinidad, satisfacción del cliente, predicción de bancarrota y cartera análisis. (párr. 2)

- **Muestra:** Muestree los datos creando una o más tablas de datos. Las muestras deben ser lo suficientemente grandes como para contener la información importante, pero lo suficientemente pequeñas como para procesarlas.
- **Explorar:** Explore los datos en busca de relaciones anticipadas, tendencias no anticipadas y anomalías para obtener información e ideas.

- **Modificar:** Modifique los datos creando, seleccionando y transformando las variables para enfocar el proceso de selección del modelo.
- **Modelo:** Modele los datos utilizando herramientas analíticas para buscar una combinación de datos que prediga de manera confiable el resultado deseado.
- **Evaluar:** Evaluar los datos mediante la evaluación de la utilidad y la fiabilidad de los resultados del proceso de minería de datos.

El proceso de minería de datos SEMMA está impulsado por un diagrama de flujo de proceso, que puede modificar y guardar. La GUI está diseñada de tal manera que el analista de negocios con poca experiencia estadística puede navegar a través de la metodología de minería de datos, mientras que el experto cuantitativo puede ir "detrás de escena" para afinar y afinar el proceso analítico.

Las fases de la metodología CRISP-DM actual para la minería de datos son las siguientes:

- **Comprensión del negocio:** Esta fase inicial se enfoca en comprender los objetivos y requisitos del proyecto desde una perspectiva comercial, y luego convertir este conocimiento en una definición de un problema de minería de datos y un plan preliminar del proyecto diseñado para lograr los objetivos.
- **Comprensión los datos:** La fase de comprensión de datos comienza con una recopilación inicial de datos y continúa con actividades para familiarizarse con los datos, identificar problemas de calidad de datos, descubrir primeras ideas sobre los datos o detectar subconjuntos interesantes para formar hipótesis para ocultar información.
- **Preparación de datos:** La fase de preparación de datos cubre todas las actividades para construir el conjunto de datos final (datos que se incorporarán a la herramienta o herramientas de modelado) a partir de los

datos sin procesar iniciales. Es probable que las tareas de preparación de datos se realicen varias veces y no en el orden prescrito.

- **Modelado:** En esta fase, se seleccionan y aplicarán diversas técnicas de modelado, y sus parámetros se calibran a valores óptimos. En general, existen varias técnicas para el mismo tipo de problema de minería de datos. Algunas técnicas requieren formatos de datos específicos.
- **Evaluación:** En esta etapa del proyecto, ha creado uno o más modelos que parecen ser de alta calidad, desde la perspectiva del análisis de datos. Antes de proceder a la implementación final del modelo, es importante evaluarlo más a fondo y revisar los pasos tomados para construirlo, a fin de garantizar que logre adecuadamente los objetivos comerciales.
- **Implementación:** La creación del modelo generalmente no es el final del proyecto. En general, el conocimiento adquirido debe organizarse y presentarse de tal manera que el cliente pueda usarlo. Dependiendo de los requisitos, la fase de implementación puede ser tan simple como generar un informe o tan compleja como implementar un proceso de minería de datos repetible. (p.6)

Knowledge Discovery in Databases (KDD), Fayyad, Shapiro y Smyth (1996) indicaron que KDD es un proceso poco común de identificación de patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles en los datos. La metodología KDD consta de las siguientes fases: selección, preprocesamiento, transformación, minería de datos y evaluación. Esto se puede resumir en los pasos que se presentarán a continuación:

- **Primer paso:** Se desarrolla una comprensión de la aplicación, debe tener un dominio y conocimiento previo.

- **Segundo paso:** Debe crear un conjunto de datos de destino: seleccionando un conjunto de datos, o centrándose en un subconjunto de variables o datos de muestra, el descubrimiento que se llevará a cabo.
- **Tercer paso:** Limpieza y preprocesamiento de datos. Las operaciones básicas incluyen eliminación de ruido, campos de datos vacíos, etc.
- **Cuarto paso:** Reducción y proyección de datos: la búsqueda de características útiles para representar los datos de acuerdo con el objetivo de la tarea (Reducción de dimensionalidad).
- **Quinto paso:** Es poner el objetivo del KDD (primer paso) a un método de minería de datos.
- **Sexto paso:** Es el análisis exploratorio y de hipótesis y el modelo de la lección secundaria: la elección del algoritmo de minería de datos que se utilizará para buscar patrones de datos.
- **Séptimo paso:** El séptimo es la minería de datos: la búsqueda de patrones de interés en una determinada forma de representación o un conjunto de tales representaciones.
- **Octavo paso:** Interpretación de los patrones minados, posiblemente puede volver a cualquiera de los pasos 1 a 7 para más iteración. Este paso también puede implicar mostrar los patrones y modelos extraídos o mostrar los datos que contienen los modelos extraídos.
- **Noveno paso:** Está actuando sobre el conocimiento descubierto: el uso del conocimiento directamente, incorporando el conocimiento en otro sistema para la adopción de nuevas medidas o, simplemente, documentación e informes a las partes interesadas. Este proceso también incluye verificar y resolver posibles conflictos con el conocimiento creído (o extraído).

Ramos (2016) explicó que el modelado dimensional se usa hoy en la mayoría de las soluciones de BI. Es una mezcla correcta de normalización y desnormalización, comúnmente llamada Normalización Dimensional. Se utiliza para el diseño de Data Marts y Data Waterhouse. Básicamente, hay dos tipos de tablas: tablas de dimensiones y tablas de hechos. (p.16) A continuación, se detallarán más de ellos:

- **Tabla de dimensiones:** Ramos (2016) detalló que una tabla de dimensiones contiene una serie de atributos o características, mediante los cuales podemos agrupar, cortar o filtrar la información. A veces, estos atributos se organizan en jerarquías que permiten analizar los datos de forma agrupada. Las tablas de dimensiones son generalmente muy anchas (contienen muchos atributos y pueden tener bastantes caracteres cada una) y cortas (generalmente tienen pocas filas). (p. 30)
- **Tabla de hechos:** Chinchilla (2011) señaló que las claves foráneas en su conjunto corresponden a la clave primaria de dicha tabla central, que, según la terminología de la base de datos multidimensional, también se denomina tabla de hechos. La tabla de hechos contiene, además de las llaves extrañas, una o más mediciones (hechos numéricos) que ocurren como resultado de combinar las dimensiones que definen cada registro. (p. 60) Por otro lado, Ramos (2016) señaló que los elementos que componen la tabla de hechos son:

(a) Clave primaria (PK): identifica de forma exclusiva cada fila. Al igual que en los sistemas transaccionales, cada tabla debe tener una clave principal, en una tabla de hechos puede tenerla o no, y esto tiene sus pros y sus contras, pero ambas posiciones son defendibles. (b) Claves foráneas (FK): señale las claves principales (claves subrogadas) de cada una de las dimensiones que están relacionadas con esa tabla de hechos y (c) Mediciones: Representan columnas que contienen datos numéricos cuantificables que se pueden agregar. Por ejemplo, Importe, cantidad, precio, margen, número de operaciones, etc. (p. 18)

Por otra parte, los especialistas de IBM (s. f.) indicaron que un esquema dimensional separa físicamente las medidas que cuantifican el negocio de los elementos descriptivos (también llamados dimensiones) que describen y categorizan el negocio. Los diagramas dimensionales son los siguientes:

- **Modelo estrella:** Alpa y Jayesh (2012) explicaron que el modelo de estrella es la estructura básica de un modelo dimensional. Tiene una mesa central grande (tabla de hechos) y un conjunto de tablas más pequeñas (dimensiones) dispuestas en un patrón radial alrededor de la mesa central (p. 243).
- **Modelo copo de nieve:** Garani y Helmer (2012) indicaron que el esquema del copo de nieve ha tratado de resolver el problema de las dependencias funcionales de transición que es evidente en un esquema en estrella al disminuir los niveles de desnormalización. Se puede obtener de un esquema en estrella eliminando progresivamente algunas dependencias funcionales transitivas existentes en las tablas de dimensiones dividiendo las tablas de dimensiones en tablas más pequeñas. Por lo tanto, el patrón de estrella se considera un caso especial del patrón de copo de nieve. (p. 26)
- **Modelo constelación:** Feki y Hachaichi (2007) argumentaron que el esquema de constelación, que es una generalización de la estrella, es preferible porque reúne dos o más hechos que comparten dimensiones comunes. Es decir, el esquema de constelación permite el análisis de hechos cruzados que examina un hecho con respecto a otros. Por lo tanto, las constelaciones intensifican los análisis, dando a los tomadores de decisiones la oportunidad de explorar y comprender mejor su desempeño comercial. (p.1)

Tabla 16: Comparación entre modelos de minería de datos

MODELO	FUNCIONALIDAD	PERSPECTIVA	TÉCNICAS UTILIZADAS
<b>Descriptivo</b>	Está diseñado para generar frecuencia y secuencia, tabulación cruzada y correlación interesante. Se pueden definir modelos descriptivos para encontrar regularidades desconocidas interesantes en los datos, para descubrir patrones ocultos y para encontrar un conjunto interesante de subgrupos en la mayor parte de los datos admitidos. (Abed y Nasr, 2019, p. 4143)	Esta perspectiva utiliza datos para responder:  <b>¿Qué pasó?</b> Esto se hará analizando eventos pasados y actuales utilizando datos históricos (Oliveira, 2016, p. 5).	El modelo descriptivo utiliza técnicas como: (a) Agrupación, (b) Reglas de asociación, (c) Resúmenes y (d) Extracción de características. Estas técnicas analizan el rendimiento pasado y entienden ese rendimiento mediante la extracción de datos históricos para encontrar las razones detrás del éxito o el fracaso pasado. (Lico, 2017, p. 44)
<b>Predictivo</b>	Está diseñado para descubrir los patrones más potentes en grandes bases de datos. El modelo clásico para los datos de predicción son los casos muestreados. El papel del modelo de minería de datos predictivo es aprender los criterios de toma de decisiones para asignar etiquetas a otros casos nuevos y no clasificados. (Ionita, 2015, p. 299)	Esta perspectiva utiliza datos para responder:  <b>¿Qué podría pasar?</b> Esto se hará extrapolando patrones o relaciones que se encuentran en datos históricos (Oliveira, 2016, p. 5).	El modelo predictivo utiliza técnicas como: (a) Clasificación, (b) Regresión, (c) Análisis de series de tiempo y (d) Detección de anomalías. Estas técnicas permiten al minero de datos predecir un valor desconocido (a menudo futuro) de una variable específica (Anita y Jasjeet, 2017, p. 25).
<b>Prescriptivo</b>	Está diseñada para proporcionar predicciones e información personalizada según el contexto. Esta técnica se utiliza para respaldar una toma de decisiones más efectiva basada en varias ideas cuando los tomadores de decisiones empresariales, como los CTO y CEO, analizan y predicen situaciones complejas. (Gim, Lee y Joo, 2018, p. 2)	Esta perspectiva utiliza datos para responder:  <b>¿Qué debemos hacer?</b> Esto se realizará mediante el análisis de datos estructurados complementados con datos semiestructurados o no estructurados, como el juicio de expertos o las preferencias de los tomadores de decisiones (Oliveira, 2016, p. 5).	El modelo prescriptivo utiliza técnicas tales como: (a) Clasificación, (b) Regresión, (c) Análisis de series de tiempo y (d) Red neuronal. Estas técnicas permiten el análisis de atributos pasados y actuales para determinar el estado futuro (Fashoto, Adeyeye, Owolabi y Odin, 2015, p. 509).

*Nota:* Según lo mencionado anteriormente, la investigación se centra en un modelo predictivo, ya que su análisis puede arrojar resultados de un futuro probable, esto ayudará a tomar buenas decisiones para enfrentar lo que se puede aproximar.

Tabla 17: Comparación entre técnicas predictivas de minería de datos

TÉCNICA	DEFINICIÓN	PROBLEMAS A SOLUCIONAR	TIPOS DE ALGORITMOS
<b>Clasificación</b>	Las técnicas de clasificación consisten en predecir un posible resultado basado en una entrada dada. La clasificación se utiliza para clasificar cada información en un grupo predefinido o conjunto de datos de clase (Ahmedy Kauser, 2017, p. 1887).	Los contextos en los que una tarea de clasificación es esencial son: <ul style="list-style-type: none"> <li>• Estado del crédito sobre la base de información financiera.</li> <li>• Diagnóstico de enfermedades de un paciente (Lico, 2017, p. 46).</li> </ul>	Los métodos comúnmente utilizados son <ul style="list-style-type: none"> <li>• Máquina de vectores de soporte (SVM)</li> <li>• Árboles de decisión</li> <li>• Red de creencias bayesianas</li> <li>• Redes neuronales (Saima, Ibrahim, Senan y N, 2018, p. 4).</li> </ul>
<b>Regresión</b>	La técnica de regresión predice valores numéricos continuos, esta técnica se utiliza para ajustar una ecuación a un conjunto de datos. Se usa con mayor frecuencia cuando el atributo de destino tiene un valor numérico. Se puede usar para modelar la relación entre una o más variables independientes y variables dependientes. (Lico, 2017, p. 46)	Los contextos en los que una tarea de regresión es esencial son: <ul style="list-style-type: none"> <li>• Volúmenes de ventas</li> <li>• Precios de las acciones</li> <li>• Tasas de falla del producto (Lico, 2017, p. 46).</li> </ul>	Los métodos comúnmente utilizados son: <ul style="list-style-type: none"> <li>• Regresión Lineal</li> <li>• Regresión no lineal (Syaidatus Syahira, Mutalib, Nurzeatul Hamimah y Shuzlina, 2019, p. 7).</li> </ul>
<b>Series de tiempo</b>	Las técnicas de series de tiempo determinan si existe algún vínculo entre los datos recopilados hasta el momento y cómo cambian los datos (Angadi y Kulkami, 2015, p. 104).	Los contextos en los que una tarea de series de tiempo es esencial es cuando surge un problema con respecto a la alta dimensionalidad de los datos y la dificultad de definir una forma de medición de similitud basada en la percepción humana. (Lico, 2017, p. 47)	Los métodos comúnmente utilizados son: <ul style="list-style-type: none"> <li>• Indexación</li> <li>• Representación</li> <li>• Medida de similitud (Chen, Deng, Wan, Zhang, Va-silakos y Rong, 2015, p. 5).</li> </ul>
<b>Detección de anomalías</b>	La detección de anomalías es el proceso de encontrar registros periféricos para un conjunto de datos dado. La detección de anomalías ha sido ampliamente estudiada en estadística y aprendizaje automático (Mohiuddin, Anwar, Abdun, Shah y Maher, 2015, p. 3).	Los contextos en los que una tarea de detección de anomalías es esencial son: <ul style="list-style-type: none"> <li>• Detección de fraude</li> <li>• Detección de sospechas de transacciones en organizaciones financieras (Lico, 2017, p. 47).</li> </ul>	Los métodos comúnmente utilizados son: <ul style="list-style-type: none"> <li>• Anomalías puntuales</li> <li>• Anomalías contextuales</li> <li>• Anomalías colectivas (Lico, 2017, p. 47).</li> </ul>

*Nota:* De acuerdo con lo anterior, se puede rescatar que el problema de investigación cubre las técnicas de regresión, ya que los valores que se deben predecir son el número de muertes y, por otro lado, las técnicas de series temporales, ayudarán a tener un análisis de pronóstico el comportamiento de los datos a lo largo del tiempo.



Tabla 18: Comparación entre metodologías de minería de datos

METODOLOGÍA	FUNCIÓN	FASES	FORTALEZAS	LIMITACIONES
<b>SEMMA</b>	Su función es llevar a cabo las tareas básicas de minería de datos y centrarse en los aspectos del desarrollo del modelo de minería de datos, utiliza el software SAS Enterprise Mine. En la metodología SEMMA, los movimientos no son estrictos durante el proyecto, puede deslizarse hacia adelante y hacia atrás y repetir los pasos. (Daderman y Rosander, 2018, p. 12)	<ul style="list-style-type: none"> <li>▪ Muestra</li> <li>▪ Explorar</li> <li>▪ Modificar</li> <li>▪ Modelo</li> <li>▪ Evaluar (Daderman y Rosander, 2018, p. 12).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Buen usuario soportado</li> <li>▪ Admiten varias técnicas de minería de datos</li> <li>▪ Proceso iterativo (Daderman y Rosander, 2018, p. 26).</li> </ul>	<ul style="list-style-type: none"> <li>• Está diseñado para funcionar con la herramienta de minería Enter-Award de SAS, surgen problemas si aparece un caso de minería no típico.</li> <li>• No considera la comprensión comercial de un problema (Daderman y Rosander, 2018, p. 26).</li> </ul>
<b>CRISP-DM</b>	Proporcionar un marco para llevar a cabo proyectos de minería de datos que sea independiente tanto del sector industrial como de la tecnología utilizada (Wirth y Hipp, 2000, p. 2).	<ul style="list-style-type: none"> <li>▪ Comprensión empresarial</li> <li>▪ Comprender los datos</li> <li>▪ Preparación de datos</li> <li>▪ Modelado</li> <li>▪ Evaluación</li> <li>▪ implementación (Wirth y Hipp, 2000, p. 5-7).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Proceso claramente definido, predicciones sobre posibles problemas.</li> <li>▪ Fase completa para transformar el problema en términos de DM.</li> <li>▪ Vale la pena para grandes proyectos</li> <li>▪ Admite varias técnicas de minería de datos</li> <li>▪ Proceso iterativo</li> <li>▪ Tiene documentación con estudios de casos típicos que ayudarán con el proceso. (Daderman and Rosander, 2018, p. 26)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Proceso largo y complicado.</li> <li>▪ La preparación de datos y las fases de modelado de los datos de transmisión son diferentes de la minería de datos estática tradicional debido a su serie de tiempo. Es posible que no esté cubierto en la documentación de crisp-dm, ya que está hecho para un enfoque general de la minería de datos. (Daderman y Rosander, 2018, p. 26)</li> </ul>
<b>KDD</b>	Identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles en los datos (Fayyad, Piatetsky-Shapiro y Smyth, 1996, p. 83).	<ul style="list-style-type: none"> <li>▪ Selección</li> <li>▪ Preprocesamiento</li> <li>▪ Transformación</li> <li>▪ Minería de Datos</li> <li>▪ Evaluación (Fayyad, Piatetsky-Shapiro y Smyth, 1996, p. 84).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Proceso iterativo</li> <li>▪ Admite varias técnicas de minería de datos, por ejemplo, redes neuronales (Daderman y Rosander, 2018, p. 26).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Requiere conocimiento previo en minería de datos (Daderman y Rosander, 2018, p. 26).</li> </ul>

*Nota:* De acuerdo con lo mencionado anteriormente, las metodologías de minería de datos ayudarán a tener un orden en el proceso de análisis de datos, para lo cual se utilizará una herramienta de criterios de puntuación, a fin de encontrar la mejor opción para la presente investigación.

Tabla 19: Selección de la metodología de minería de datos

	FUNCIÓN		FASES		FORTALEZAS		DEBILIDADES		TOTAL
METODOLOGÍA	20 %		30 %		35 %		15 %		100 %
SEMA	1.00	<b>0.20</b>	3.00	<b>0.90</b>	1.00	<b>0.35</b>	2.00	<b>0.30</b>	<b>1.75</b>
CRISP-DM	3.00	<b>0.60</b>	2.00	<b>0.60</b>	3.00	<b>1.05</b>	3.00	<b>0.45</b>	<b>2.70</b>
KDD	2.00	<b>0.40</b>	2.00	<b>0.60</b>	2.00	<b>0.70</b>	1.00	<b>0.15</b>	<b>1.85</b>

*Nota:* De lo que se detalla anteriormente, se puede ver que los criterios de evaluación fueron: función, fases, fortalezas y limitaciones, que se evaluarán con una puntuación de 1 a 3 y cada criterio tendrá un peso porcentual, esta herramienta multiplica el puntaje por el peso del criterio de evaluación, para cada criterio habrá un puntaje total, al final se agregará y el ganador será el que tenga el puntaje más alto, para la presente investigación se decidió utilizar la metodología CRISP- DM, porque salió con un puntaje total de 2.70, siendo superior a las otras metodologías de minería de datos.

## Anexo 4: SISTEMA DE INFORMACIÓN GEOGRÁFICA

### SISTEMA DE INFORMACIÓN GEOGRÁFICA

Foote y Lynch (1986) explicaron:

La importancia del GIS como tecnología integradora también es evidente en su pedigrí. El desarrollo de GIS se ha basado en innovaciones realizadas en muchas disciplinas diferentes: geografía, cartografía, fotogrametría, teledetección, topografía, geodesia, ingeniería civil, estadística, informática, investigación operativa, inteligencia artificial, demografía y muchas otras ramas de la sociedad. Las ciencias naturales y la ingeniería han contribuido. De hecho, algunas de las aplicaciones más interesantes de la tecnología SIG que se analizan a continuación se basan en este carácter y patrimonio interdisciplinario. (p. 1)

Las herramientas GIS con mayor presencia en el mercado tecnológico se enumeran a continuación:

- **ArcGIS:** ArcGIS es una herramienta utilizada para modelar una ubicación o un mapa. Al respecto, Kumar (2017) explicó que ArcGIS Server y ArcGIS Desktop, para construir un SIG completo, ArcGIS Desktop es el marco básico para compilar, mapear, editar, analizar y visualizar datos espaciales. Es un conjunto integrado de aplicaciones de escritorio, como ArcMap, Arc Catalog, Arc Toolbox y Arc Globe. (p. 1) Con el tiempo, la tecnología ha evolucionado, y es por eso que el desarrollo o diseño de un mapa hoy en día es más fácil gracias a la tecnología ArcGIS.

Según Kumar (2017) indicó que, en los últimos años, se ha generado una gran cantidad de datos como resultado de las aplicaciones informáticas automatizadas existentes. El avance en las redes de computadoras y el desarrollo de aplicaciones distribuidas, como el comercio electrónico y las redes sociales, ha dado paso a la generación de grandes conjuntos de datos. Es difícil procesar un volumen tan grande

de datos con formas secuenciales tradicionales. El gran volumen de datos ha llevado a la adopción de procesamiento paralelo para el procesamiento eficiente de datos en un conjunto de máquinas informáticas colaboradas. La parte MapReduce de Hadoop abstrae todos los detalles del procesamiento paralelo del usuario y el usuario obtiene un marco muy simplificado para la programación. El marco consiste en componentes de mapeador y reductor que funcionan en el concepto de pares de valores clave. (p. 1)

- **QGIS:** Khan y Aaqib (2017) señalaron que QGIS (anteriormente conocido como Quantum GIS) es una aplicación de sistema de información geográfica de escritorio (GIS) multiplataforma libre y de código abierto que admite la visualización, edición y análisis de datos geoespaciales. Por lo tanto, Gary Sherman comenzó el desarrollo de Quantum GIS a principios de 2002, y se convirtió en un proyecto de incubadora de Open Source Geospatial Foundation en 2007. La versión 1.0 fue lanzada en enero de 2009. Como una aplicación de software libre bajo GNU GPL, QGIS puede modificarse libremente para realizar tareas diferentes o más especializadas. (p. 727-728)
- **GRASS GIS:** GRASS GIS, comúnmente conocido como GRASS (Sistema de Soporte de Análisis de Recursos Geográficos), es un paquete de software gratuito y de código abierto del Sistema de Información Geográfica (SIG) utilizado para la gestión y análisis de datos geoespaciales, procesamiento de imágenes, producción de gráficos y mapas, modelado espacial y visualización. Las raíces del proyecto GRASS GIS surgieron por primera vez hace 33 años. GRASS GIS se utiliza actualmente en entornos académicos y comerciales de todo el mundo, así como en muchas agencias gubernamentales y empresas de consultoría ambiental. Es miembro fundador de la Open Source Geospatial Foundation (OSGeo). (Khan y Aaqib, 2017, p. 728)
- **gvSIG Desktop:** Fue fundado por el Consejo Regional de Infraestructuras y Transporte (CIT) de Valencia (España) para reemplazar el software de

funcionalidad similar al ArcView de ESRI en las autoridades municipales, ya que el gobierno regional tiene como objetivo cambiar todos los sistemas a una computadora basada en Linux infraestructura. El desarrollo de gvSIG había comenzado a fines de 2003 y está liderado por la empresa IVER S.A. (España). Varias universidades y otras compañías están incluidas en el proyecto, gvSIG es conocido por su interfaz fácil de usar y por poder acceder a todos los formatos vectoriales y ráster. (Khan y Aaqib, 2017, p. 728)

- **SAGA:** SAGA (Sistema de Análisis Geocientíficos Automatizados) es un sistema de información geográfica (SIG) de código abierto utilizado para editar y analizar datos espaciales. Incluye una gran cantidad de módulos para el análisis de datos vectoriales (punto, línea y polígono), tabla y cuadrícula e imagen. Entre otros, el paquete incluye módulos para geoestadística, clasificación de imágenes, proyecciones, simulación de procesos dinámicos (hidrología, desarrollo del paisaje) y análisis del terreno. Se puede acceder a la funcionalidad a través de una GUI, la línea de comando o mediante la API de C ++. (Khan y Aaqib, 2017, p. 728)

Tabla 20: Comparación entre herramientas GIS

SOFTWARE	MULTI-PLATAFORMA	DESCRIPCIÓN	HERRAMIENTAS	PLATAFORMA SOPORTADA	EXTENSIÓN DE ARCHIVO	CANTIDAD DE LENGUAJES	COSTO
<b>ARCGIS</b>	SI	Le permite crear aplicaciones creadas en mapas y resolver problemas con análisis espacial.	Para topografía. calles. terrenos y océanos. Capa de datos operativa, simbología, etiquetado y rango de escala.	Windows 7 o superior	DGN, DXF, DWG, SHP	11 lenguajes.	Licencia Comercial.
<b>QGIS</b>	SI	Ofrece 400 complementos de complementos. Diseñado para la capacidad 3D de mapeo ping.	Herramientas de calidad. extensiones de plugins para geografía y análisis espacial.	Linux, Unix. Windows, Mac OSX y Android	69 formatos de vectores. 100 formatos de rastreadores	40 lenguajes.	Gratis.
<b>GRASS GIS</b>	SI	Gestión de tierras y planificación ambiental. No líder en mapeo.	Procesamiento digital de terreno y estadísticas, 350 herramientas de manipulación trama y vector.	Windows. Linux y sitios.	ESRI, SHP, DXF	11 lenguajes.	Gratis.
<b>GVSIG DESKTOP</b>	SI	Supera a QGIS en opciones 3D. Alta compatibilidad. Buenas herramientas CAD. Tiene un manejo de datos ordenado.	Permite trazar geometrías, alinear, editar vértices, dividir líneas y polígonos.	Windows y Linux.	ESRI, SHP, DXF	20 lenguajes.	Gratis.
<b>SAGA</b>	SI	Cuenta con herramientas geocientíficas.	Análisis de terreno. extracción de cuenca. sombreado y análisis de visibilidad, morfométrico, índice de humedad.	Windows, Linux, Mac OS X y GNU.	ESRI, SHP, DXF	10 lenguajes.	Gratis

Nota: Adaptado de "Evaluación empírica de ArcGIS con soluciones contemporáneas de código abierto: un estudio" (De Khan y Aaqib, 2017, Revista internacional de ingeniería e investigación tecnológica).

Tabla 21: Selección de la herramienta GIS

SOFTWARE	HERRAMIENTAS		PLATAFORMA SOPORTADA		EXTENSIÓN DE ARCHIVO		INTERFACES USADOS POR EL LENGUAJE		COSTO		TOTAL
	3.00	0.70	3.00	0.75	3.00	0.45	2.00	0.20	1.00	0.15	2.60
ARCGIS	3.00	<b>1.05</b>	3.00	<b>0.75</b>	3.00	<b>0.45</b>	2.00	<b>0.20</b>	1.00	<b>0.15</b>	<b>2.60</b>
QGIS	2.00	<b>0.70</b>	3.00	<b>0.75</b>	1.00	<b>0.15</b>	3.00	<b>0.30</b>	3.00	<b>0.45</b>	<b>2.35</b>
GRASS GIS	2.00	<b>0.70</b>	2.00	<b>0.50</b>	2.00	<b>0.30</b>	2.00	<b>0.20</b>	3.00	<b>0.45</b>	<b>2.15</b>
GVGIS DESKTOP	2.00	<b>0.70</b>	2.00	<b>0.50</b>	1.00	<b>0.15</b>	2.00	<b>0.20</b>	3.00	<b>0.45</b>	<b>2.00</b>
SAGA	2.00	<b>0.70</b>	2.00	<b>0.50</b>	2.00	<b>0.30</b>	2.00	<b>0.20</b>	3.00	<b>0.45</b>	<b>2.15</b>

Nota: Por las razones explicadas anteriormente, se realizó una evaluación basada en los siguientes criterios: herramientas, plataforma admitida, extensión de archivo, idiomas para interfaces de usuario y costo, en los cuales la herramienta ganadora fue ArcGIS con una puntuación de 2.60.

## **Anexo 5: CLASIFICACIÓN INTERNACIONAL DE ENFERMEDADES**

### CLASIFICACIÓN INTERNACIONAL DE ENFERMEDADES

La Clasificación Internacional de Enfermedades juega un papel muy importante en esta investigación, ya que cada uno de sus grandes grupos se utilizó como variables predictoras, además de mantener un orden internacional de enfermedades. Por lo tanto, los especialistas de la OMS (s. f.) explicaron:

La primera edición de la clasificación internacional, conocida como la Lista Internacional de Causas de Muerte, fue adoptada por el Instituto Internacional de Estadística en 1893. A la OMS se le confió el ICD en su creación en 1948 y publicó la sexta versión. CIE-6, que incorporaba la morbilidad por primera vez. El Reglamento de Nomenclatura de la OMS, adoptado en 1967, estipuló que los Estados Miembros utilizan la revisión más reciente de la CIE para las estadísticas de mortalidad y morbilidad. El ICD ha sido revisado y publicado en una serie de ediciones para reflejar los avances en salud y ciencias médicas a lo largo del tiempo. La CIE-10 fue aprobada en mayo de 1990 por la cuadragésima tercera Asamblea Mundial de la Salud. Se cita en más de 20,000 artículos científicos y es utilizado por más de 100 países alrededor del mundo.

La jerarquía más alta en la CIE-10 son los capítulos de enfermedades, que se clasificarán de la siguiente manera:

- 1. A00-B99:** Ciertas enfermedades infecciosas y parasitarias.
- 2. C00-D48:** Tumores [neoplasias]
- 3. D50-D89:** Enfermedades de los órganos formadores de sangre y ciertos trastornos relacionados con el mecanismo inmune.
- 4. E00-E90:** Enfermedades endocrinas, nutricionales y metabólicas.
- 5. F00-F99:** Trastornos mentales y del comportamiento.



- 6. G00-G99:** Enfermedades del sistema nervioso
- 7. H00-H59:** Enfermedades del ojo y anexos.
- 8. H60-H95:** Enfermedades del oído y del proceso mastoideo.
- 9. I00-I99:** Enfermedades del sistema circulatorio
- 10. J00-J99:** Enfermedades del sistema respiratorio
- 11. K00-K93:** Enfermedades del sistema digestivo
- 12. L90-L99:** Enfermedades de la piel y del tejido subcutáneo.
- 13. M00-M99:** Enfermedades del sistema músculo esquelético y del tejido conectivo.
- 14. N00-N99:** Enfermedades del sistema genitourinario.
- 15. O00-O99:** Embarazo, parto y puerperio.
- 16. P00-P96:** Ciertas condiciones que se originan en el período perinatal.
- 17. Q00-Q99:** Malformaciones congénitas, deformaciones y anomalías cromosómicas.
- 18. R00-R99:** Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte
- 19. S00-T98:** Lesiones, intoxicaciones y otras consecuencias de causas externas.
- 20. V01-Y98:** Causas externas de morbilidad y mortalidad.

Cabe señalar que lo anterior pertenece a la jerarquía más alta de la CIE-10, porque esta clasificación se divide en 4 jerarquías que se mencionarán en orden de mayor a menor: (a) capítulos, (b) grupos, (c) categorías y (d) enfermedades.

## **Anexo 6: Desarrollo de la metodología CRISP-DM**

### **DESARROLLO DE LA METODOLOGÍA CRISP-DM**

En la presente investigación se realizará el proceso de minería de datos con la metodología CRISP-DM siguiendo cada una de sus fases. A continuación, se presentará el desarrollo de cada etapa de la metodología:

#### **1. COMPRENSIÓN DEL NEGOCIO**

En esta sección, se seguirá cada una de las tareas de esta primera fase de la metodología CRISP – DM, cuyo propósito es determinar los objetivos y requisitos del proyecto desde una perspectiva de negocio, para posteriormente poder convertirlos en objetivos desde el punto de vista técnico y terminar con un plan de proyecto donde se indicara los tiempos e inconvenientes que pueden surgir.

##### **1.1 Determinar los objetivos de negocio**

El objetivo en este proyecto es hacer predicciones tan completas como sea posible a partir de los datos que ya están disponibles sobre las muertes registradas en Chile. El objetivo es proporcionar un conocimiento útil, que permita una buena toma de decisiones o la planificación de nuevas investigaciones de dichas fuentes, a través de una aplicación web.

- **Contexto:** El pronóstico relacionado con la mortalidad se realiza en base a información veraz compartida por una organización. En referencia a la situación actual del negocio en la organización, existe una base de datos de defunciones obtenidas por el Departamento de Estadística e Información de Salud (DEIS), desde la cual se puede obtener información sobre las personas fallecidas, como también la causa de la muerte. Sin embargo, no se ha encontrado un estudio en profundidad sobre el comportamiento de la mortalidad a partir del cual se puedan sacar

conclusiones o patrones para hacer predicciones sobre la cantidad de defunciones futuras.

- **Objetivos de Negocio:** Los objetivos de negocio, como ya se mencionó, el objetivo general es proporcionar un conocimiento útil a través de la predicción de muertes, lo que permitirá una buena toma de decisiones o nuevas investigaciones de estas fuentes. Los objetivos de este proyecto se han definido de la siguiente manera:
  - Encontrar un modelo que explique la mortalidad de Chile.
  - Pronosticar la cantidad de defunciones para un futuro cercano.

La información obtenida pueden ser de mucha utilidad para los especialistas de la salud para plantear nuevas estrategias, así como para detectar las enfermedades más problemáticas en cada región, provincia o comuna de Chile y tratar de averiguar por qué ciertas enfermedades están presentes en mayor cantidad en comparación con otros lugares, ya sea debido a la falta de preparación de los servicios de salud, la falta de conciencia de la población, etc. Todo esto permitirá a los especialistas y autoridades de salud mejorar su toma de decisiones.

- **Criterios de éxito de negocio:** La posibilidad de hacer predicciones sobre la cantidad de defunciones con un alto porcentaje de exactitud se establece como un criterio de éxito, de modo que se pueda brindar información útil sobre el comportamiento de la mortalidad en Chile. Por ejemplo, encontrar la enfermedad que tiene una mayor impacto en el incremento de la mortalidad en cada región. Al proporcionar esta información, las autoridades podrían plantear nuevas estrategias y tomar decisiones certeras.

## 1.2 Valoración de la situación

Se cuenta con una base de datos con información detallada sobre las personas que fallecieron entre los años 1990 y 2017, por lo que se puede afirmar que hay

información más que suficiente para analizar y aplicar técnicas de minería de datos. Esta información incluye la causa de muerte, la base que respalda el diagnóstico, el lugar de ocurrencia y otros datos del fallecido que pueden ser útiles en el proceso de minería de datos.

## Inventario de Recursos

En esta sección, se mostrará la situación actual de los recursos con lo cual se cuenta para la realización del proceso de minería de datos.

- **Datos:** Se cuenta con la base de datos de defunciones publicada por la DEIS, el cual contiene información detallada de las defunciones de Chile desde el año 1990 a 2017.
- **Hardware:** Los recursos de hardware disponibles son 2 computadoras portátiles con las siguientes características:

Tabla 22: *Comprensión del negocio - Inventario de recursos a nivel de hardware*

Características	Laptop 1	Laptop 2
<b>Marca</b>	TOSHIBA	TOSHIBA
<b>Modelo</b>	Satellite C55-C	Satellite C55-C
<b>Procesador</b>	I5 - 5ta CPU @ 2.20GHz	I3 - 5ta CPU @ 2.20GHz
<b>RAM</b>	4GB	4GB
<b>Almacenamiento</b>	500GB	500GB
<b>Tarjeta Gráfica</b>	Intel HD Graphics 5500	Intel HD Family
<b>Sistema operativo</b>	Windows 10	Windows 8.1

- **Software:** Con respecto a los recursos de software, una herramienta de minería de datos y un sistema de gestión de bases de datos (DBMS) están disponibles porque el origen de la fuente de datos se encontró en formato .csv, por lo que tuvo que usarse esta herramienta para poder importar los datos para un mejor control. Estos programas tienen las siguientes características:

Tabla 23: *Comprensión del negocio - Inventario de recursos a nivel de software*

<b>Características</b>	<b>DBMS</b>	<b>DM AND ML</b>
<b>Nombre</b>	SQL Server	Microsoft Azure ML
<b>Versión</b>	2019 Express	2020
<b>Desarrollado en</b>	C++	-
<b>Desarrollador por</b>	Microsoft	Microsoft
<b>Sistema Operativo</b>	Microsoft. Linux y Docker	Nube

### **Requisitos, supuestos y contingencias.**

Los requisitos para el proyecto fueron:

- Se requiere tener la información de los códigos únicos territoriales de cada de región de Chile, esta información se podrá obtener desde la Subsecretaría de Desarrollo Regional y Administrativo (SUBDERE).
- Se requiere tener la información de los códigos de CIE-10, debido a que es un estándar internacionalizado impulsado por la OMS para el intercambio de información entre las entidades de la salud de todo el mundo.
- Se requiere normalizar la información, para posteriormente cargarla a un motor de base de datos para el análisis y el proceso de minería.

El supuesto principal para el proyecto fue que los resultados deberían corresponder a la tasa de mortalidad esperada en el período de tiempo analizado

en la predicción. Por otro lado, no se encontraron restricciones porque los datos están disponibles públicamente en la página oficial de la DEIS.

### **Costo y beneficios**

Los datos de este proyecto no suponen ningún costo adicional, debido a que como se mencionó anteriormente, el acceso a los datos es público. En cuanto a los beneficios, no se puede decir que este proyecto genere ningún beneficio económico para el sector de la salud y las autoridades directamente, pero se puede suponer indirectamente, ya que el objetivo de este proyecto es proporcionar un conocimiento útil, lo que permite la buena toma de decisiones o el enfoque de nuevas investigaciones de dichas fuentes.

### **1.3. Determinar los Objetivos de la Minería de Datos**

#### **Metas de minería de datos**

Los objetivos en términos de minería de datos son:

- Determinar el impacto de la técnica de regresión lineal múltiple en el pronóstico de defunciones.
- Determinar el impacto de la técnica de regresión logística ordinal en el pronóstico de defunciones
- Determinar el impacto de la técnica de series de tiempo en el pronóstico de defunciones

#### **Criterios de éxito de minería de datos**

Desde el punto de vista de la minería de datos, la posibilidad de hacer predicciones sobre la cantidad muertes en el futuro con un alto porcentaje de precisión se establece como un criterio de éxito, específicamente se definirá este porcentaje al menos en un 80% con la técnica de regresión lineal múltiple, un

60% con la técnica de regresión logística ordinal y un MAPE inferior al 5% con la técnica de series de tiempo, todo esto sustentando en las hipótesis planteadas al inicio de la investigación. El grado de precisión estará definido por cada técnica que se utilizará al obtener el modelo de minería de datos, por lo que este tema se abordará nuevamente más adelante en el paso 5 (evaluación) de la metodología.

#### 1.4. Producir el plan de proyecto

El proyecto se dividirá en las siguientes etapas para tener un orden y estimar el tiempo de finalización

Tabla 24: *Comprensión del negocio - Plan del proyecto*

<b>Etapas</b>	<b>Descripción</b>	<b>Duración</b>	<b>Riesgo</b>
<b>Etapas 1</b>	Análisis de la estructura de los datos e información en la base de datos.	2 semanas	Poca comprensión del contexto
<b>Etapas 2</b>	Ejecución de consultas para obtener muestras representativas de los datos.	1 semana	Problemas de datos, problemas de TI.
<b>Etapas 3</b>	Preparación de datos (selección, limpieza, conversión y formateo, si es necesario) para facilitar la extracción de datos en ellos.	1 semana	Elegir datos apropiados
<b>Etapas 4</b>	Elección de técnicas de modelado y ejecución de las mismas en los datos.	2 semanas	Dificultad para encontrar el modelo correcto, problemas tecnológicos.
<b>Etapas 5</b>	Análisis de los resultados obtenidos en la etapa anterior, si fuera necesario repetir la etapa 4.	2 semanas	Dificultad para evaluar resultados ambiguos
<b>Etapas 6</b>	Producción de informes con los resultados obtenidos en función de los objetivos del negocio y los criterios de éxito establecidos.	1 semana	Dificultad en el informe final fácil de entender
<b>Etapas 7</b>	Presentación de los resultados finales.	1 semana	Resultados desfavorables

## **2. COMPRESIÓN DE LOS DATOS**

En esta segunda fase de la metodología CRISP-DM, la recopilación de datos inicial se realiza para establecer un primer contacto con el problema, familiarizarse con los datos y descubrir su calidad, así como identificar las relaciones más obvias para formular las primeras hipótesis.

### **2.1. Recolectar los datos iniciales**

Los datos utilizados en este proyecto esta basada en la información de fallecidos de Chile que se encuentra en el portal web de la DEIS, por otro lado se utilizará una base de datos externa de los códigos únicos territoriales obtenidos por la SUBDERE y para la codificación estándar de las enfermedades se contara con la CIE-10 publicada por la OMS. Cabe mencionar que las bases de datos que se obtuvieron de dichos portales web se encontraron en formato CSV y XLS, por lo que toda la información tuvo que cargarse en un administrador de bases de datos.

En este caso se usó SQL Server para un mejor control de los datos. La base de datos de defunciones de la DEIS está compuesta por 97 atributos, de los cuales 34 son las glosas de las columnas. Por lo tanto, se realizó una separación de las glosas para insertarlas por separado en las tablas creadas previamente. A continuación se muestra una imagen del proceso general de recopilación e importación de datos en SQL Server.



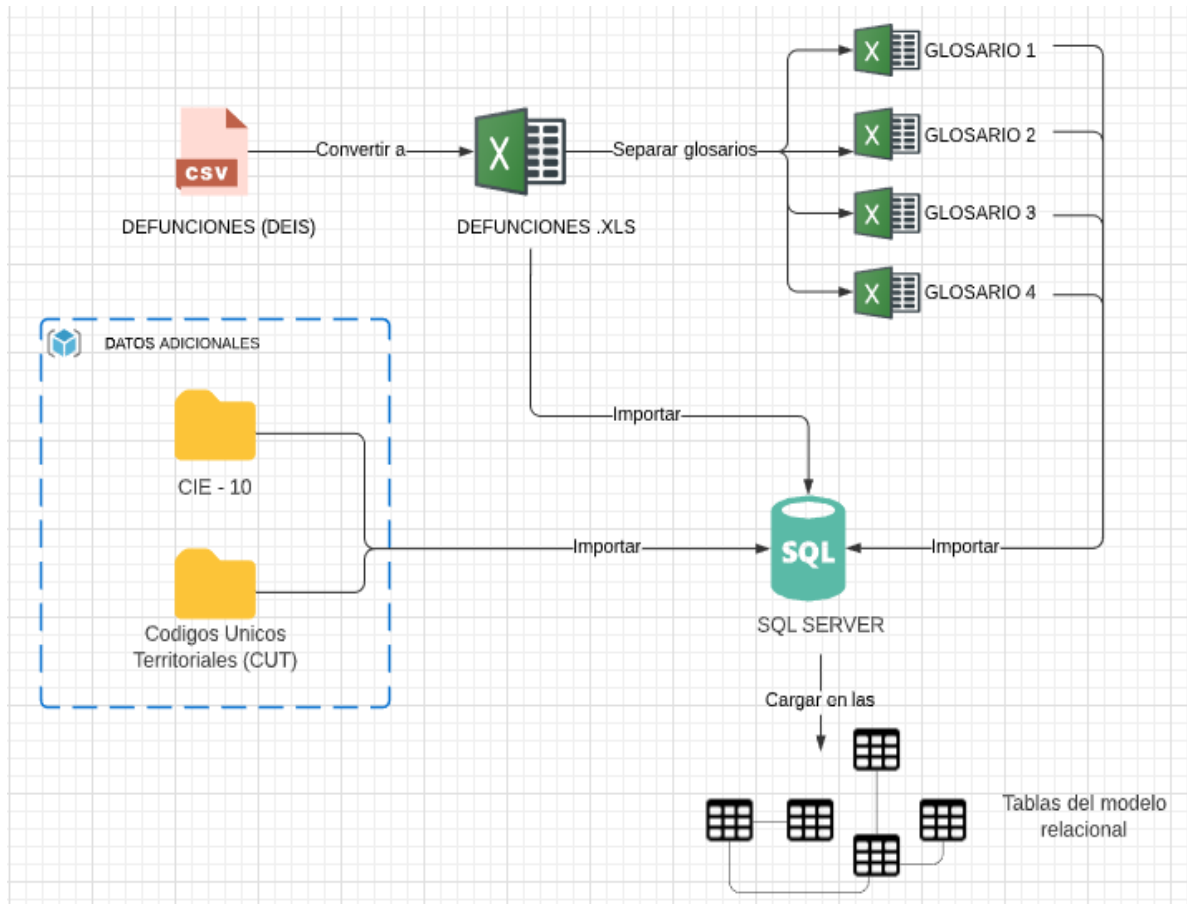


Figura 10: Comprensión de los datos - Recolección de datos iniciales

## 2.2 Descripción de los datos

En esta etapa se describirá cada atributo que conforma la base de datos de las defunciones, por lo cual a continuación se presentara el diagrama del modelo relacional.

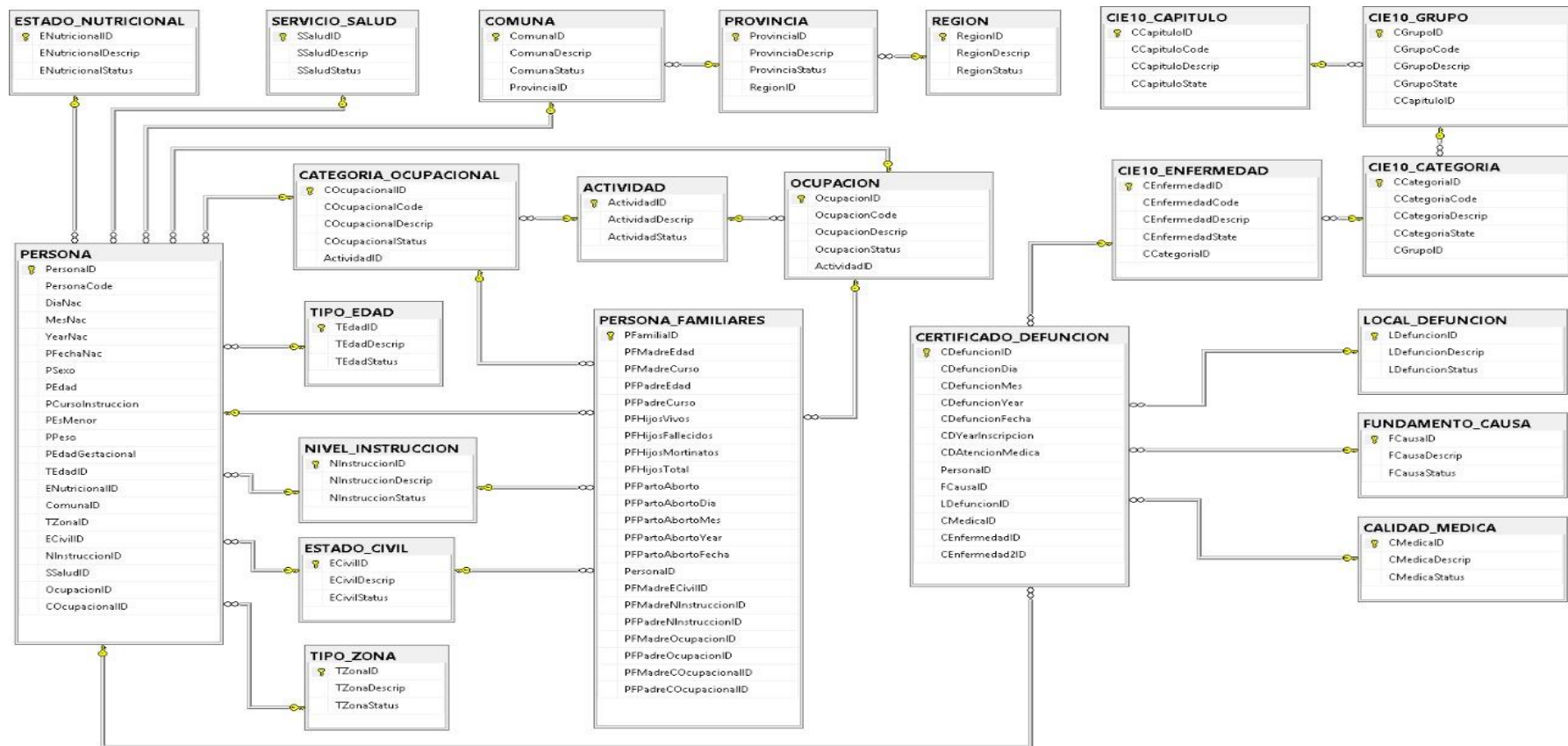


Figura 11: Comprensión de los datos - Diagrama de la base de datos relacional

A continuación, se explicará cada tabla del modelo relacional, además de sus atributos, la descripción de cada atributo y el tipo de datos.

Tabla 25: *Comprensión de los datos - Tabla ESTADO\_CIVIL*

Columna	Tipo de datos	Descripción	Valores
<b>EcivilID</b>	INT	Este campo representa el código del estado civil del difunto.	1: Soltero 2: Casado 3: Viudo 4: Divorciado
<b>ECivilDescrip</b>	VARCHAR	Este campo representa el brillo del estado civil.	5: Separado Judicial 6: Conviviente Civil 9: Ignorado

*Nota: Esta tabla tendrá el estado civil de la persona fallecida: su clave principal es EcivilID. Esta tabla tiene un total de 7 registros.*

Tabla 26: *Comprensión de los datos - Tabla TIPO\_EDAD*

Columna	Tipo de datos	Descripción	Valores
<b>TEdadID</b>	INT	Es el código de la unidad que se mide a la edad.	1: Edad en años 2: Edad en meses 3: Edad en días 4: Edad en horas
<b>TEdadDescrip</b>	VARCHAR	Es el glosario o descripción de la edad.	0: Ignorado

*Nota: Esta tabla indica la unidad de medida para la edad: su clave principal es TEdadID. Esta tabla tiene un total de 6 registros.*

Tabla 27: *Comprensión de los datos - Tabla CALIDAD\_MEDICA*

Columna	Tipo de datos	Descripción	Valores
<b>CMedicalID</b>	INT	Código que identifica la calidad de quien certifica la muerte	1: Médico tratante 2: Médico Legista - Patólogo 3: Otro 4: Información de matronas
<b>CMedicaDescrip</b>	VARCHAR	Código que identifica la calidad de quien certifica la muerte	5: Testigo 9: Ignorado

*Nota: Esta tabla indica la calidad de la persona que certifica la muerte: su clave principal es CMedicalID. Esta tabla tiene un total de 6 registros.*

Tabla 28: *Comprensión de los datos - Tabla NIVEL\_INSTRUCCION*

Columna	Tipo de datos	Descripción	Valores
<b>NInstruccionID</b>	INT	Código de nivel educativo del fallecido: indica el nivel educativo del fallecido como se indica en el certificado médico de defunción.	1: Superior 2: Medio 3: Secundaria 4: Básico o Primario 5: Ninguno 9: Ignorado
<b>NInstruccionDescrip</b>	VARCHAR	Descripción del nivel de educación.	

*Nota:* Esta tabla indica el nivel de educación del fallecido, indica el nivel educativo del fallecido según lo indicado por el médico de la muerte: Su clave principal es NInstruccionID. Esta tabla tiene un total de 6 registros.

Tabla 29: *Comprensión de los datos - Tabla TIPO\_ZONA*

Columna	Tipo de datos	Descripción	Valores
<b>TZonalID</b>	INT	Código para definir el área urbana o rural de la residencia de la persona fallecida	1: Urbano 2: Rural
<b>TZonaDescrip</b>	VARCHAR	Glosario o descripción de la zona	

*Nota:* Esta tabla indica la definición del área urbana o rural de la residencia de la persona fallecida: su clave principal es TZonalID. Esta tabla tiene un total de 2 registros.

Tabla 30: *Comprensión de los datos - Tabla LOCAL\_DEFUNCION*

Columna	Tipo de datos	Descripción	Valores
<b>LDefuncionID</b>	INT	Código del lugar de ocurrencia de la muerte.	1: Hospital o Clínica 2: Casa Habitación 3: Otro 9: Ignorado
<b>LDefuncionDescrip</b>	VARCHAR	Describe dónde ocurre la muerte.	

*Nota:* Esta tabla indica el lugar de fallecimiento de la persona: su clave principal es LDefuncionID. Esta tabla tiene un total de 4 registros.

Tabla 31: *Comprensión de los datos - Tabla FUNDAMENTO\_CAUSA*

Columna	Tipo de datos	Descripción	Valores
<b>FCausalID</b>	INT	Código de motivos para la causa de la muerte	1: Autopsia 2: Biopsia 3: Operación 4: Ex. Laboratorio
<b>FCausaDescrip</b>	VARCHAR	Glosario de la causa de la muerte	5: Cuadro Clínico 6: Información de Testigos

*Nota:* Esta tabla indica la causa raíz de la muerte: su clave principal es FCausalID. Esta tabla tiene un total de 6 registros.

Tabla 32: *Comprensión de los datos - Tabla ESTADO\_NUTRICIONAL*

Columna	Tipo de datos	Descripción	Valores
<b>ENutricionalID</b>	INT	Código de estado nutricional antes de la enfermedad cuando se trata de la muerte de niños menores de un año de edad.	1: Eutrófico 2: Desnutrición grado I 3: Desnutrición grado II 4: Desnutrición grado III 9: Ignorado
<b>ENutricionalDescrip</b>	VARCHAR	Glosario estado nutricional cuando se trata de la muerte de menos de un año.	

*Nota:* Esta tabla indica el estado nutricional cuando se trata de la muerte de menos de un año: su clave principal es ENutricionalID. Esta tabla tiene un total de 5 registros

Tabla 33: *Comprensión de los datos - Tabla REGION*

Columna	Tipo de datos	Descripción	Valores
<b>RegionID</b>	INT	Código de la región de residencia del difunto.	Los valores están basados e los codigos únicos territoriales (CUT) publicados por la SUBDERE. Chile se divide en 16 Regiones
<b>RegionDescrip</b>	VARCHAR	Brillo de la región de residencia.	

*Nota:* Esta tabla indica la región de residencia del difunto: su clave principal es RegionID. Esta tabla tiene un total de 16 registros.

Tabla 34: *Comprensión de los datos - Tabla PROVINCIA*

Columna	Tipo de datos	Descripción	Valores
<b>ProvincialID</b>	INT	Código de la provincia de residencia del difunto	Los valores están basados e los codigos únicos territoriales (CUT) publicados por la SUBDERE, Chile se divide en 57 provincias
<b>ProvinciaDescrip</b>	VARCHAR	Brillo de la provincia de residencia del difunto	
<b>RegionID</b>	INT	Clave secundaria de la tabla REGION.	

*Nota:* Esta tabla indica la región de residencia del difunto: su clave principal es ProvincialID. Esta tabla tiene un total de 57 registros

Tabla 35: *Comprensión de los datos - Tabla COMUNA*

Columna	Tipo de datos	Descripción	Valores
<b>ComunaD</b>	INT	Código de la comuna de residencia del difunto	Los valores están basados e los codigos únicos territoriales (CUT) publicados por la SUBDERE. Chile se divide en 347 comunas
<b>ComunaDescrip</b>	VARCHAR	Brillo de la comuna de residencia	
<b>ProvincialID</b>	INT	Brillo de la comuna de residencia	

*Nota:* Esta tabla indica la región de residencia del difunto: su clave principal es ComunaID. Esta tabla tiene un total de 347 registros

Tabla 36: *Comprensión de los datos - Tabla SERVICIO\_SALUD*

Columna	Tipo de datos	Descripción	Valores
<b>SsaludID</b>	INT	Servicio de Salud código de residencia del difunto	Codificación de acuerdo con el Sistema Nacional de Servicios de Salud (SNSS) que tiene 29 Servicios de salud territoriales que tienen responsabilidades para la acción de salud en territorios geográficos definidos, que pueden ser de alcance regional o subregional
<b>SsaludDescrip</b>	VARCHAR	Glosario del servicio de salud de la residencia	

*Nota:* Esta tabla indica el servicio de salud de residencia del difunto: su clave principal es SsaludID. Esta tabla tiene un total de 29 registros.

Tabla 37: *Comprensión de los datos - Tabla ACTIVIDAD*

Columna	Tipo de datos	Descripción	Valores
<b>ActividadID</b>	INT	Código de condición de actividad	0: Inactivo 1: Actividad 2: Cesante o Desocupado 9: Ignorado
<b>ActividadDescrip</b>	VARCHAR	Glosario de códigos de actividad	

*Nota:* Esta tabla indica el servicio de salud de residencia del difunto: su clave principal es ActividadID. Esta tabla tiene un total de 4 registros.

Tabla 38: *Comprensión de los datos - Tabla CATEGORIA\_OCUPACIONAL*

Columna	Tipo de datos	Descripción	Valores
<b>COcupacionalID</b>	INT	Id de la categoría ocupacional	0: Inactivo 1: Patrón 2: Empleado 3: Obrero 4: Trabajador por cuenta propia 9: Ignorado
<b>COcupacionalCode</b>	INT	Código condicionado a los datos de la actividad.	
<b>COcupacionalDescrip</b>	VARCHAR	Brillo del código de categoría ocupacional	

*Nota:* Esta tabla indica la categoría ocupacional: su clave principal es COcupacionalID y su clave externa es ActividadID. Esta tabla tiene un total de 6 registros.

Tabla 39: *Comprensión de los datos - Tabla OCUPACION*

Columna	Tipo de datos	Descripción	Valores
<b>OcupacionID</b>	INT	Identificación de la ocupación de la persona fallecida.	
<b>OcupacionCode</b>	INT	Código condicionado a los datos de la actividad.	Depende de la variable ACTIVIDAD. La ocupación se basa en los grandes grupos de actividad de la Clasificación Internacional Uniforme de Ocupaciones (CIUO)
<b>OcupacionalDescrip</b>	VARCHAR	Glosario de Ocupación	
<b>ActividadID</b>	INT	Clave secundaria de la tabla ACTIVIDAD	

*Nota:* Esta tabla indica cuál era la ocupación del difunto: su clave principal es OcupacionID y su clave secundaria es ActividadID. Esta tabla tiene un total de 19 registros.

Tabla 40: *Comprensión de los datos - Tabla PERSONA*

<b>Columna</b>	<b>Tipo de datos</b>	<b>Descripción</b>
<b>PersonalID</b>	INT	Identificación de la persona fallecida
<b>PersonaCode</b>	INT	Identificador único y anónimo de la persona fallecida.
<b>DiaNac</b>	INT	Fecha de nacimiento
<b>MesNac</b>	INT	Mes de fecha de nacimiento
<b>YearNac</b>	INT	Año de fecha de nacimiento
<b>PFechaNac</b>	DATE	en formato DD / MM / AA
<b>PSexo</b>	INT	Código que identifica el sexo biológico
<b>PEdad</b>	INT	Cantidad de edad: es el número de la edad.
<b>PCursoInstruccion</b>	INT	Último curso aprobado por la persona fallecida.
<b>PEsMenor</b>	INT	Código para identificar si el fallecido es menor de 1 año.
<b>PPeso</b>	INT	Peso al nacer en gramos
<b>PEdadGestacional</b>	INT	Edad gestacional en semanas
<b>TEdadID</b>	INT	Clave secundaria de la tabla TIPO_EDAD
<b>ENutricionalID</b>	INT	Clave secundaria de la tabla ESTADO_NUTRICIONAL
<b>ComunalID</b>	INT	Clave secundaria de la tabla COMUNA
<b>TZonalID</b>	INT	Clave secundaria de la tabla TIPO_ZONA
<b>ECivilID</b>	INT	Clave secundaria de la tabla ESTADO_CIVIL
<b>NInstruccionID</b>	INT	Clave secundaria de la tabla NIVEL_INSTRUCCION
<b>SSaludID</b>	INT	Clave secundaria de la mesa SERVICIO_SALUD
<b>OcupacionID</b>	INT	Clave secundaria de la tabla OCUPACION
<b>COcupacionalID</b>	INT	Clave secundaria de la tabla CATEGORIA_OCUPACIONAL

*Nota: Esta tabla indica los datos personales del fallecido: su clave principal es PersonalID.*



Tabla 41: *Comprensión de los datos - Tabla PERSONA\_FAMILIARES*

Columna	Tipo de datos	Descripción
<b>PFamilialID</b>	INT	Identificación del pariente de la persona fallecida.
<b>PFMadreEdad</b>	INT	Edad de la madre en años cuando se trata de la muerte del niño menor de un año.
<b>PFMadreCurso</b>	INT	Último curso de instrucción de la madre cuando se trata de la muerte del niño menor de un año.
<b>PFPadreEdad</b>	INT	Edad del padre en años cuando se trata de la muerte de un niño menor de un año.
<b>PFPadreCurso</b>	INT	El último curso de instrucción del padre cuando se trata de la muerte de un niño menor de un año.
<b>PFHijosVivos</b>	INT	Número de niños vivos cuando se trata de la muerte de menos de un año.
<b>PFHijosFallecidos</b>	INT	Número de niños fallecidos cuando se trata de la muerte de un niño menor de un año.
<b>PFHijosMortinatos</b>	INT	Número de niños nacidos muertos.
<b>PFHijosTotal</b>	INT	Número total de niños cuando se trata de la muerte de menos de un año.
<b>PFPartoAborto</b>	INT	Parto o aborto.
<b>PFPartoAbortoDia</b>	INT	Día del último parto o aborto cuando se trata de la muerte de menos de un año.
<b>PFPartoAbortoMes</b>	INT	Mes del último parto o aborto cuando se trata de la muerte de menos de un año.
<b>PFPartoAbortoYear</b>	INT	Año del último parto o aborto cuando se trata de la muerte de un niño menor de un año.
<b>PersonalID</b>	INT	Clave secundaria de la tabla PERSONA.
<b>PFMadreEcivilID</b>	INT	Clave secundaria de la tabla ESTADO_CIVIL.
<b>PFMadreNInstruccionID</b>	INT	Clave secundaria de la tabla NIVEL_INSTRUCCION.
<b>PFPadreNInstruccionID</b>	INT	Clave secundaria de la tabla NIVEL_INSTRUCCION.
<b>PFMadreOcupacionID</b>	INT	Clave secundaria de la tabla OCCUPACION.
<b>PFPadreOcupacionID</b>	INT	Clave secundaria de la tabla OCCUPACION.
<b>PFMadreCOcupacionalID</b>	INT	Clave secundaria de la tabla CATEGORIA_OCUPACIONAL.
<b>PFPadreCOcupacionalID</b>	INT	Clave secundaria de la tabla CATEGORIA_OCUPACIONAL.

*Nota: Esta tabla indica los datos personales de la familia del difunto.*

Tabla 42: *Comprensión de los datos - Tabla CERTIFICADO\_DEFUNCION*

Columna	Tipo de datos	Descripción
<b>CDefuncionID</b>	INT	ID del certificado de defunción
<b>CDefuncionDia</b>	INT	Día de certificación médica
<b>CDefuncionMes</b>	INT	Mes de certificación medica
<b>CDefuncionYear</b>	INT	Año de certificación médica
<b>CDefuncionFecha</b>	DATE	Fecha de Certificación médica
<b>CDYearInscripcion</b>	INT	Fecha de registro del año de defunción
<b>CDAtencionMedica</b>	INT	Código para identificar si recibió atención médica
<b>PersonalID</b>	INT	Clave secundaria de la tabla PERSONA
<b>FCausalID</b>	INT	Clave secundaria de la tabla FUNDAMEN-TO_CAUSA
<b>LDefuncionID</b>	INT	Clave secundaria de la tabla LOCAL_DEFUNCION
<b>CMedicalID</b>	INT	Clave secundaria de la tabla CALIDAD_MEDICA
<b>CEnfermedadID</b>	INT	Clave secundaria de la mesa CIE10_ENFERMEDAD
<b>CEnfermedad2ID</b>	INT	Clave secundaria de la mesa CIE10_ENFERMEDAD

*Nota:* Esta tabla indica el certificado de defunción de la persona: su clave principal es CDefuncionID y su clave secundaria (clave externa) es PersonalID, FCausalID, LDefuncionID, CMedicalID, CEnfermedadID, CEnfermedad2ID.

Tabla 43: *Comprensión de los datos - Tabla CIE10\_CAPITULO*

Columna	Tipo de datos	Descripción	Valores
<b>CCapituloID</b>	INT	ID del capítulo de enfermedades	Los valores se basan en la codificación estándar CIE-10
<b>CCapituloCode</b>	INT	Código CIE de capítulo de enfermedad	
<b>CCapituloDescrip</b>	VARCHAR	Descripciones de capítulos de enfermedades.	

*Nota:* Esta tabla indica los capítulos de la Clasificación Internacional de Enfermedades (CIE): su clave principal es CCapituloID, esta tabla tiene un total de 20 registros.

Tabla 44: *Comprensión de los datos - Tabla CIE10\_GRUPO*

Columna	Tipo de datos	Descripción	Valores
<b>CGrupolD</b>	INT	ID del grupo de enfermedad	Los valores se basan en la codificación estándar CIE-10
<b>CGrupoCode</b>	INT	Código CIE del grupo de enfermedad	
<b>CGrupoDescrip</b>	VARCHAR	Descripciones de grupos de enfermedades.	
<b>CCapituloID</b>	INT	Clave de tabla secundaria CIE10_CAPITULO	

*Nota:* Esta tabla indica los capítulos de la Clasificación Internacional de Enfermedades (CIE): su clave principal es CGrupolD, su clave secundaria es CCapituloID

Tabla 45: *Comprensión de los datos - Tabla CIE10\_CATEGORIA*

Columna	Tipo de datos	Descripción	Valores
<b>CCategorialID</b>	INT	Id de las categorías de enfermedades	Los valores se basan en la codificación estándar CIE-10
<b>CCategoriaCode</b>	INT	Código CIE de la categoría de enfermedad	
<b>CCategoriaDescrip</b>	VARCHAR	Descripciones de categorías de enfermedades	
<b>CGrupolD</b>	INT	Tabla secundaria CIE10_GRUPO	

*Nota:* Esta tabla indica los capítulos de la Clasificación Internacional de Enfermedades (CIE): su clave principal es CCategorialID, su clave secundaria es CGrupolD.

Tabla 46: *Comprensión de los datos - Tabla CIE10\_ENFERMEDAD*

Columna	Tipo de datos	Descripción	Valores
<b>CEnfermedadID</b>	INT	Identificación de enfermedades	Los valores se basan en la codificación estándar CIE-10
<b>CCategoriaCode</b>	INT	Código CIE de la enfermedad	
<b>CCategoriaDescrip</b>	VARCHAR	Descripciones de la enfermedad	
<b>CCategorialID</b>	INT	Tabla secundaria CIE 10 CATEGORIA	

*Nota:* Esta tabla indica las Enfermedades de la Clasificación Internacional de Enfermedades (CIE): su clave principal es CEnfermedadID, su clave secundaria es CCategorialID.

## 2.3. Exploración de los datos

Una vez que se han descrito los datos, se procederá a explorarlos, esto implica aplicar pruebas estadísticas básicas que revelarán las propiedades de los datos mediante gráficos. Este informe servirá principalmente para determinar la consistencia e integridad de los datos. A través de consultas SQL, se han obtenido los datos necesarios para hacer los gráficos que se explican a continuación.

```
SELECT RegionDescrip AS 'REGION', [1] AS 'MASCULINO', [2] AS 'FEMENINO', [9] AS 'INDETERMINADO' FROM
(
SELECT PE.PersonaID,REG.RegionDescrip,PE.PSexo
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
WHERE REG.RegionID NOT IN (99)
) P PIVOT
(
COUNT (PersonaID) FOR PSexo IN ( [1], [2], [9] )
) AS pvt
ORDER BY pvt.RegionDescrip;
```

Figura 12: Consulta en la base de datos de muertes por sexo y región.

Tabla 47: *Cantidad de muertes por sexo y región*

REGIÓN	SEXO		
	Masculino	Femenino	Indefinido
Aisén del Gral. C. Ibáñez del Campo	6256	3718	0
Antofagasta	30822	24300	6
Arica y Parinacota	12280	9640	1
Atacama	15654	11786	0
Biobío	96972	80215	6
Coquimbo	39867	33601	0
La Araucanía	66617	54774	14
Libertador General Bernardo O Higgins	56346	43880	2
Los Lagos	54472	42211	6
Los Ríos	28312	21809	1
Magallanes y de La Antártica Chilena	11735	8710	0
Maule	69450	54026	6
Metropolitana de Santiago	367275	348830	45
Ñuble	35703	28113	6
Tarapacá	13927	10663	3
Valparaíso	116675	107496	6
<b>TOTAL</b>	<b>1022363</b>	<b>883772</b>	<b>102</b>

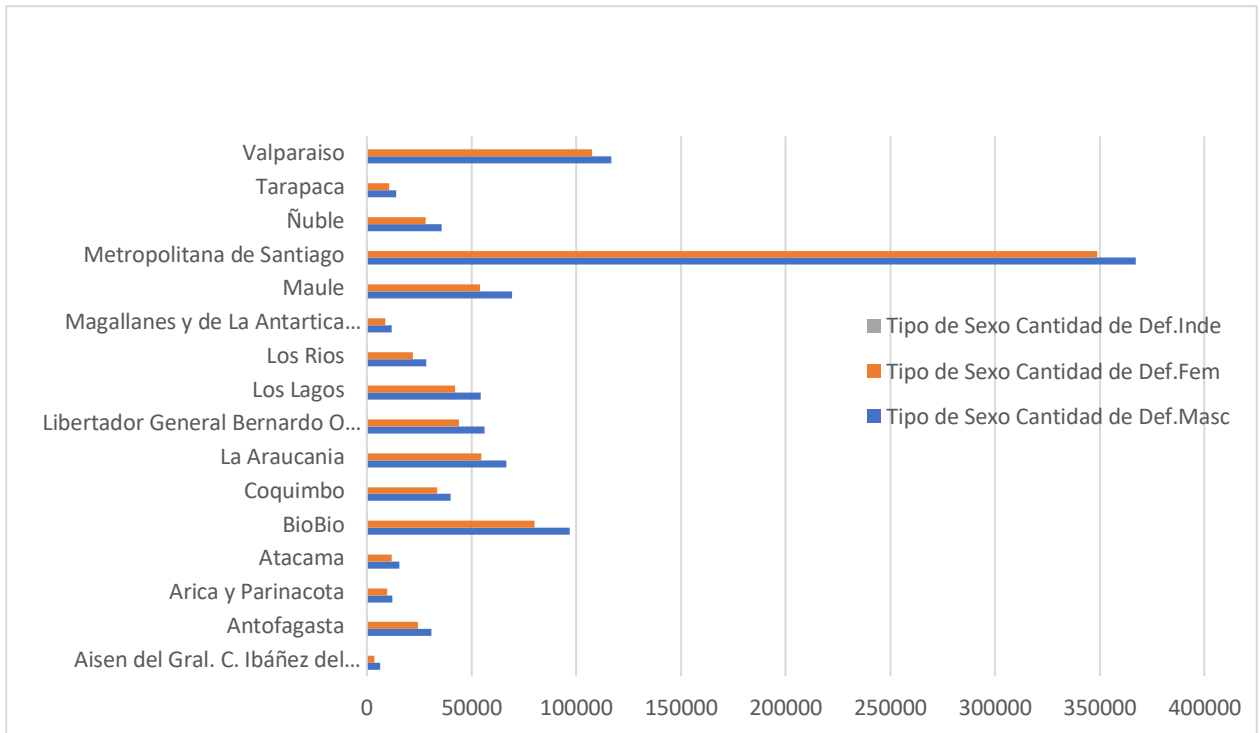


Figura 13: Cantidad de muertes por sexo y región.

*Nota:* Se puede ver que la región con más muertes es Santiago, siendo el sexo masculino con más muertes, superando las 350 mil.

```

SELECT RegionDescrip AS 'REGION', [1] AS 'SOLTERO', [2] AS 'CASADO', [3] AS 'VIUDO',
[4] AS 'DIVORCIADO', [5] AS 'SEPARADO JUDICIAL', [6] AS 'CONVIVIENTE CIVIL' FROM
(
SELECT PE.PersonaID, REG.RegionDescrip, PE.EcivilID
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
INNER JOIN ESTADO_CIVIL EC ON EC.EcivilID=PE.EcivilID
WHERE REG.RegionID NOT IN (99)
) P PIVOT
(
COUNT (PersonaID) FOR ECivilID IN ( [1], [2], [3],[4], [5], [6] )
) AS pvt
ORDER BY pvt.RegionDescrip;

```

Figure 14: Consulta en la base de datos de muertes por estado civil y región.

Tabla 48: Cantidad de muertes por estado civil y región

REGIÓN	ESTADO CIVIL					
	Soltero	Casado	Viudo	Divorciado	Separado Judicial	Conviviente Civil
Aisén del Gral. C. Ibáñez del Campo	4695	4609	569	49	0	0
Antofagasta	21863	28650	4087	232	1	2
Arica y Parinacota	8130	11806	1731	115	0	1
Atacama	12061	13337	1843	93	0	1
Biobío	62730	97236	15702	684	5	4
Coquimbo	28676	37982	6149	289	1	2
La Araucanía	46869	63019	10704	388	3	2
Libertador General Bernardo O Higgins	35899	54517	8898	387	2	3
Los Lagos	37444	49766	8789	329	3	2
Los Ríos	19342	26156	4263	154	1	2
Magallanes y de La Antártica Chilena	7256	10979	2022	90	0	1
Maule	45426	66881	10146	474	2	9
Metropolitana de Santiago	223556	413604	72751	2797	29	47
Ñuble	24106	33403	5845	197	6	1
Tarapacá	9555	12862	1856	141	0	1
Valparaíso	75100	124817	22110	960	8	20
<b>Total</b>	<b>662708</b>	<b>1049624</b>	<b>177465</b>	<b>7379</b>	<b>61</b>	<b>98</b>

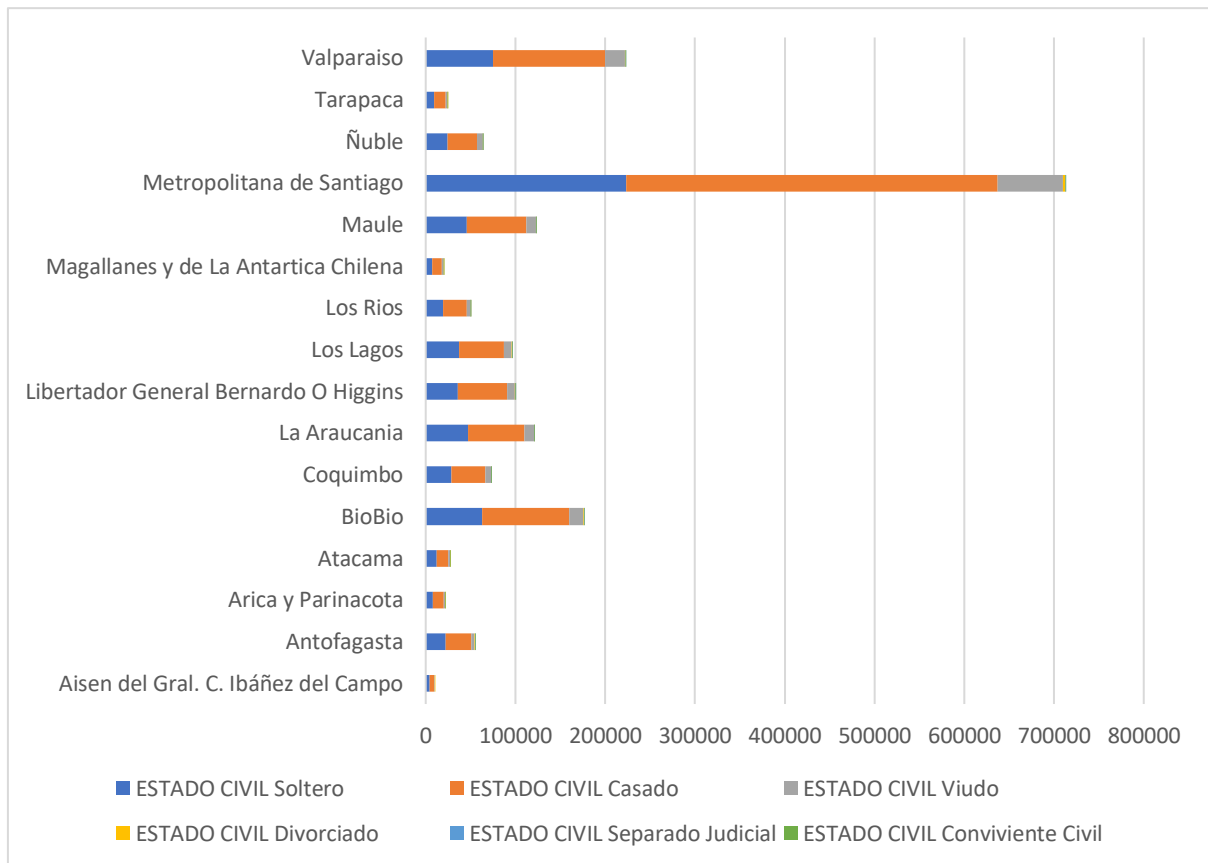


Figura 15: Cantidad de muertes por estado civil y región.

*Nota:* Se puede observar la distribución de los estados civiles: Soltero, Casado, Viudo, Divorciado, Separado Judicial y Civil sobreviviente de muertes en las Regiones. Donde se encuentra la tasa de mortalidad más alta en Metropolitana Santiago con estado civil Casado.

```

SELECT RegionDescrip AS 'REGION', [1] AS 'URBANO',
[2] AS 'RURAL' FROM
(
SELECT PE.PersonaID,REG.RegionDescrip,PE.TZonaID
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
INNER JOIN TIPO_ZONA TP ON TP.TZonaID=PE.TZonaID
WHERE REG.RegionID NOT IN (99)
) P PIVOT
(
COUNT (PersonaID) FOR TZonaID IN ( [1], [2] )
) AS pvt
ORDER BY pvt.RegionDescrip;

```

Figura 16: Consulta en la base de datos de cantidad de muertes por tipo de zona y región.

Tabla 49: Cantidad de muertes por tipo de zona y región

REGIÓN	Tipo de Zona	
	Urbano	Rural
Aisén del Gral. C. Ibáñez del Campo	7934	3718
Antofagasta	54328	24300
Arica y Parinacota	20967	9640
Atacama	25628	11786
Biobío	151420	80215
Coquimbo	58173	33601
La Araucanía	78274	54774
Libertador General Bernardo O Higgins	70066	43880
Los Lagos	61673	42211
Los Ríos	32681	21809
Magallanes y de La Antártica Chilena	19831	8710
Maule	85873	54026
Metropolitana de Santiago	696832	348830
Nuble	40257	28113
Tarapacá	23828	10663
Valparaíso	208775	107496
<b>Total</b>	<b>1636540</b>	<b>883772</b>

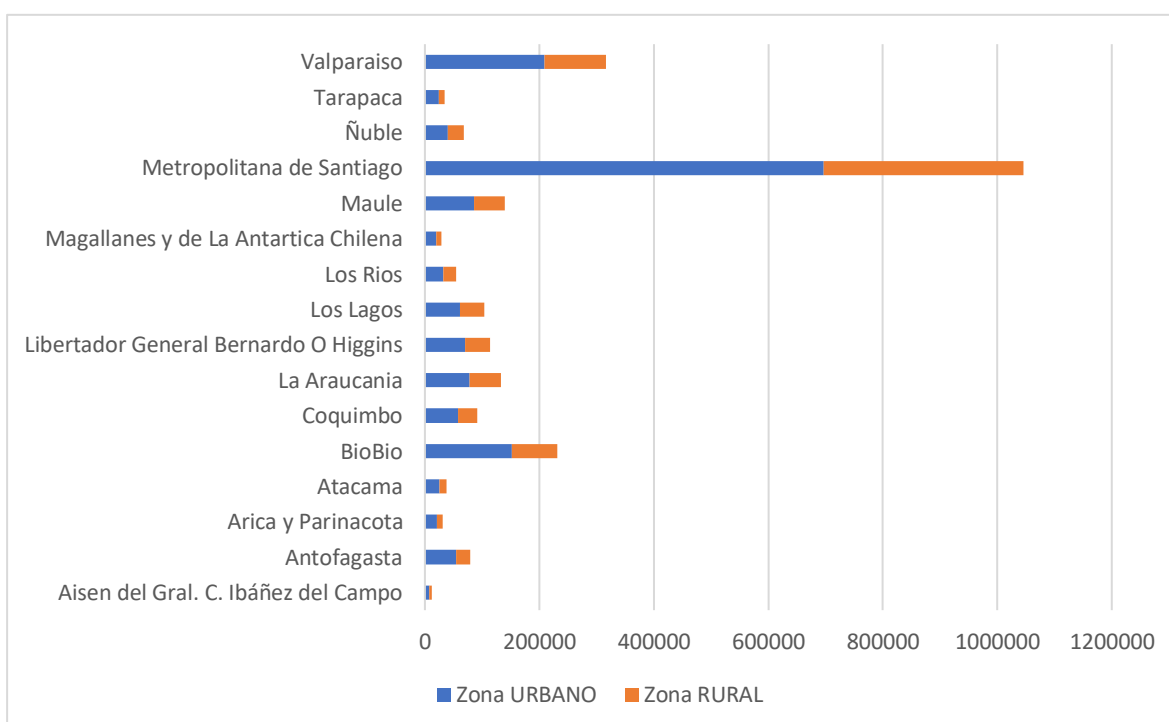


Figura 17: Cantidad de muertes por tipo de zona y región.

*Nota:* Se muestra la distribución de las zonas: Urbana y Rural de defunciones en las Regiones. Donde la tasa de mortalidad más alta es del área urbana en la región de Santiago.



```

SELECT RegionDescrip as 'REGION', [0] AS 'INACTIVO'
,[1] AS 'ACTIVO',[2] AS 'CESANTE O DESOCUPADO',[9] AS 'IGNORADO' FROM
(
SELECT PE.PersonaID,REG.RegionDescrip,A.ActividadID
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
INNER JOIN TIPO_ZONA TP ON TP.TZonaID=PE.TZonaID
INNER JOIN OCUPACION O ON O.OcupacionID=PE.OcupacionID
INNER JOIN ACTIVIDAD A ON A.ActividadID=O.ActividadID
WHERE REG.RegionID NOT IN (99)
) P PIVOT
(
COUNT (PersonaID) FOR ActividadID IN ( [0], [1], [2], [9] )
) AS pvt
ORDER BY pvt.RegionDescrip;

```

Figura 18: Consulta en la base de datos de cantidad de muertes por actividad y región.

Tabla 50: Cantidad de muertes por actividad y región

REGIÓN	ACTIVIDAD			
	Inactivo	Activo	Cesante o Desocupado	Ignorado
Aisén del Gral. C. Ibáñez del Campo	7130	2705	112	27
Antofagasta	43844	10537	618	129
Arica y Parinacota	17497	4186	197	41
Atacama	22816	4425	148	51
Biobío	150116	23636	3093	348
Coquimbo	62975	9980	403	110
La Araucanía	101365	18292	1519	229
Libertador General Bernardo O Higgins	84244	14677	1062	245
Los Lagos	78458	16714	1331	186
Los Ríos	42015	7046	990	71
Magallanes y de La Antártica Chilena	16619	3580	207	39
Maule	101936	20023	1254	269
Metropolitana de Santiago	603200	97381	13903	1666
Ñuble	53242	9580	840	160
Tarapacá	18729	5423	345	96
Valparaíso	196496	25131	2127	423
<b>Total</b>	<b>1600682</b>	<b>273316</b>	<b>28149</b>	<b>4090</b>

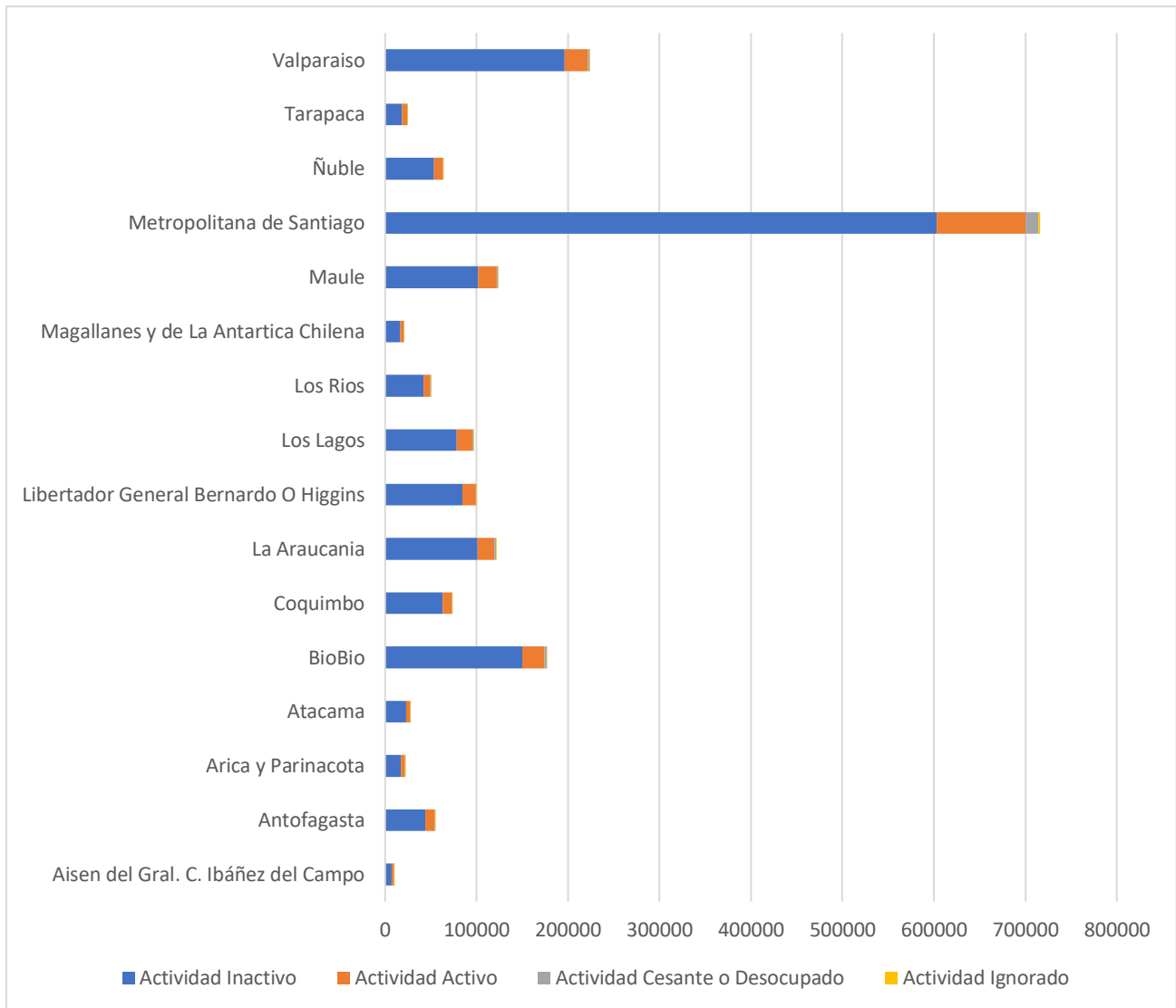


Figura 19: Cantidad de muertes por actividad y región.

*Nota:* Se muestra la distribución de las Actividades: Inactivo, Activo, Desempleado o Desempleado e Ignorado del número de muertes en las Regiones. Donde la tasa de mortalidad más alta se encuentra en Metropolitana Santiago con el nombre de la actividad inactiva.

```

SELECT RegionDescrip AS 'REGION',
[1] AS 'Ciertas enfermedades infecciosas y parasitarias',[2] AS 'Tumores [Neoplasias]',
[3] AS 'Enfermedades de la sangre y de los órganos hematopoyéticos, y ciertos trastornos que afectan el mecanismo de la inmunidad ',
[4] AS 'Enfermedades endocrinas, nutricionales y metabólicas',
[5] AS 'Trastornos mentales y del comportamiento',[6] AS 'Enfermedades del sistema nervioso central',
[7] AS 'Enfermedades del ojo y sus anexos ',[8] AS 'Enfermedades del oído y de la apófisis mastoides ',
[9] AS 'Enfermedades del sistema circulatorio ',[10] AS 'Enfermedades del sistema respiratorio',
[11] AS 'Enfermedades del sistema digestivo',[12] AS 'Enfermedades de la piel y del tejido subcutáneo',
[13] AS 'Enfermedades del sistema osteomuscular y del tejido conjuntivo',[14] AS 'Enfermedades del sistema genitourinario',
[15] AS 'Embarazo, parto y puerperio',[16] AS 'Ciertas afecciones originadas en el período perinatal',
[17] AS 'Malformaciones congénitas, deformidades y anomalías cromosómicas',
[18] AS 'Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte',
[19] AS 'Traumatismos, envenenamientos y algunas otras consecuencias de causas externas',[20] AS 'Causas externas de morbilidad y de mortalidad' FROM
(
SELECT PE.PersonaID,REG.RegionDescrip,CCAP.CCapituloID
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
INNER JOIN CERTIFICADO_DEFUNCION CD ON CD.PersonaID=PE.PersonaID
INNER JOIN CIE10_ENFERMEDAD EN ON CD.CEnfermedadID=EN.CEnfermedadID or CD.CEnfermedad2ID=EN.CEnfermedadID
INNER JOIN CIE10_CATEGORIA CATE ON EN.CCategoriaID=CATE.CCategoriaID
INNER JOIN CIE10_GRUPO GRUE ON GRUE.CGrupoID=CATE.CGrupoID
INNER JOIN CIE10_CAPITULO CCAP ON CCAP.CCapituloID=GRUE.CCapituloID
WHERE REG.RegionID NOT IN (99)
) P PIVOT
(
COUNT (PersonaID) FOR CCapituloID IN ( [1], [2], [3], [4], [5], [6], [7], [8], [9] , [10], [11],[12],[13],[14],[15],[16],[17],[18],[19],[20] )
) AS pvt

```

Figura 20: Consulta en la base de datos de cantidad de muertes por capítulo CIE-10 y región.

Tabla 51: Cantidad de muertes por capítulo de CIE-10 y región

REGION	Capítulo de Enfermedades																			
	Ciertas enfermedades infecciosas y parasitarias	Tumores [Neoplasias]	Enfermedades de la sangre y de los órganos hematopoyéticos, y ciertos trastornos que afectan el mecanismo de la inmunidad	Enfermedades endocrinas, nutricionales y metabólicas	Trastornos mentales y del comportamiento	Enfermedades del sistema nervioso central	Enfermedades del ojo y sus anexos	Enfermedades del oído y de la apófisis mastoides	Enfermedades del sistema circulatorio	Enfermedades del sistema respiratorio	Enfermedades del sistema digestivo	Enfermedades de la piel y del tejido subcutáneo	Enfermedades del sistema osteomuscular y del tejido conjuntivo	Enfermedades del sistema genitourinario	Embarazo, parto y puerperio	Ciertas afecciones originadas en el período perinatal	'Malformaciones congénitas, deformidades y anomalías cromosómicas	Síntomas, signos y hallazgos anormales y de laboratorio, no clasificados en otra parte	Traumatismos, envenenamientos y algunas otras consecuencias de causas externas	Causas externas de morbilidad y de mortalidad
Aisen del Gral. C. Ibáñez del Camp	212	2532	41	312	183	219	0	0	2398	1008	667	41	64	255	10	123	114	328	1467	1467
Antofagasta	1328	15565	265	2653	1486	1385	0	4	13369	4968	4015	202	318	1808	52	787	729	1348	4846	4846
Arica y Parinacota	784	5651	97	1381	400	664	0	1	4955	1908	1814	59	175	783	24	219	288	599	2119	2119
Atacama	675	6658	125	1331	421	725	2	1	7162	3173	1793	83	168	799	15	462	425	935	2487	2487
BioBio	4130	45014	752	7846	3477	5104	4	17	48004	17113	15293	654	882	4805	84	1831	1808	3762	16613	16613
Coquimbo	2198	19504	330	3337	1068	2017	1	3	19872	7372	4614	254	437	2011	63	771	846	3320	5450	5450
La Araucanía	2666	28849	486	4823	2480	3200	4	3	30526	12483	7440	477	625	3191	69	1054	1166	9472	12391	12391
Libertador General Bernardo O Higg	2008	22994	484	4733	2060	2634	2	6	26898	11574	7697	433	479	2873	72	955	976	3496	9854	9854
Los Lagos	2240	22853	420	4222	1820	2610	3	4	23897	10138	7135	407	551	2495	41	935	1025	5308	10585	10585
Los Ríos	1153	12046	182	2117	1035	1433	2	1	13707	5214	3188	233	276	1360	29	388	466	2313	4979	4979
Magallanes y de La Antártica Chilena	437	5136	95	908	524	525	0	2	6060	1752	1659	90	121	629	8	174	186	284	1855	1855
Maule	2437	29983	572	5678	2341	3081	1	8	35298	13493	8214	423	590	3378	55	1047	1178	3197	12508	12508
Metropolitana de Santiago	16172	178680	2700	38191	19084	22518	10	37	200013	74731	53997	1863	3978	19383	356	6697	7796	14681	55263	55263
Ñuble	1413	14847	357	3022	1438	1552	1	6	18627	6007	5423	276	291	1691	29	423	572	1563	6284	6284
Tarapaca	806	5951	78	1184	693	793	0	2	6142	1959	1948	88	174	782	22	321	351	669	2630	2630
Valparaíso	4859	56472	1047	10226	5324	6621	5	6	68618	22068	15349	1012	1196	8009	105	1546	1970	5025	14719	14719
<b>Total</b>	<b>43518</b>	<b>472735</b>	<b>8031</b>	<b>91964</b>	<b>43834</b>	<b>55081</b>	<b>35</b>	<b>101</b>	<b>525546</b>	<b>194961</b>	<b>140246</b>	<b>6595</b>	<b>10325</b>	<b>54252</b>	<b>1034</b>	<b>17733</b>	<b>19896</b>	<b>56300</b>	<b>164050</b>	<b>164050</b>

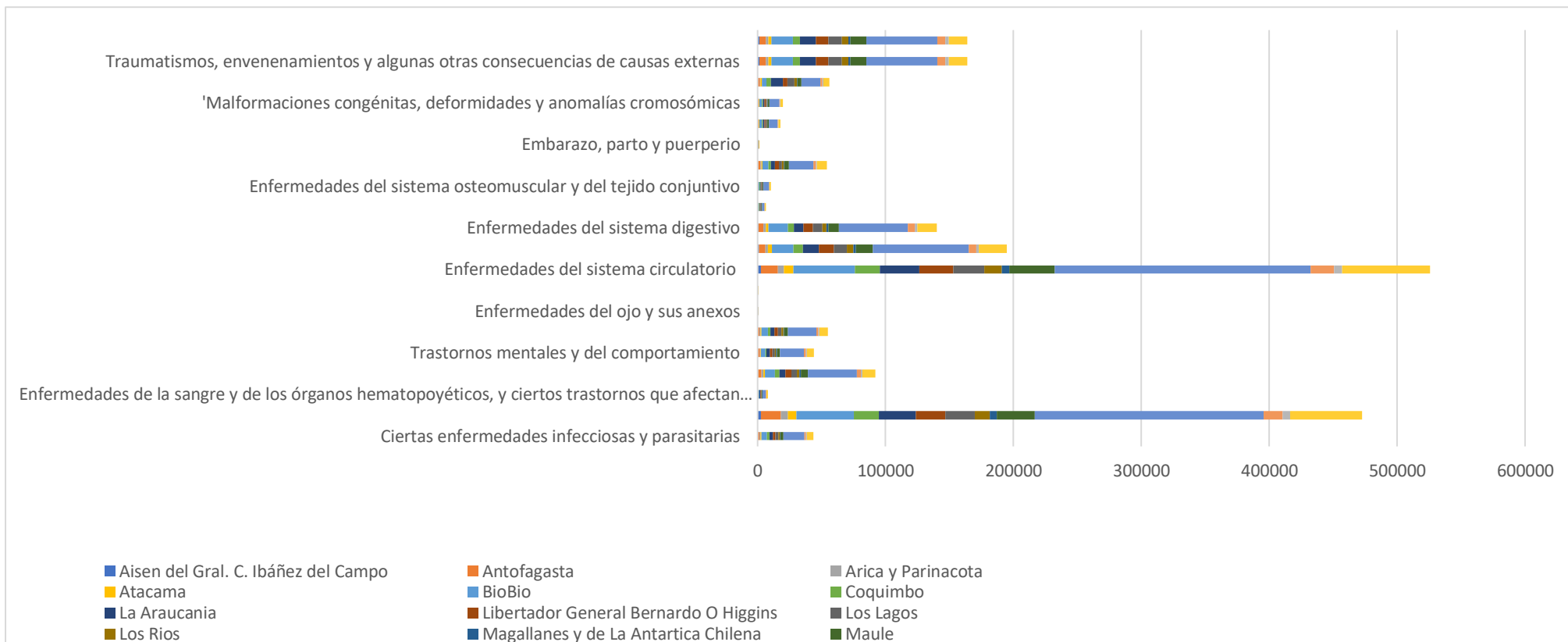


Figura 21: Cantidad de muertes por capítulo CIE-10 y región.

Nota: Se muestra la distribución de capítulos de enfermedades por número de muertes en regiones. Donde la tasa de mortalidad más alta se encuentra en la Metropolitana de Santiago con el capítulo de enfermedad del Sistema Circulatorio.

```

SELECT RegionDescrip AS 'REGION'
, [1997] AS '[1997]', [1998] AS '[1998]', [1999] AS '[1999]', [2000] AS '[2000]', [2001] AS '[2001]', [2002] AS '[2002]', [2003] AS '[2003]',
[2004] AS '[2004]', [2005] AS '[2005]', [2006] AS '[2006]', [2007] AS '[2007]', [2008] AS '[2008]', [2009] AS '[2009]', [2010] AS '[2010]',
[2011] AS '[2011]', [2012] AS '[2012]', [2013] AS '[2013]', [2014] AS '[2014]', [2015] AS '[2015]', [2016] AS '[2016]', [2017] AS '[2017]' FROM
(
SELECT PE.PersonaID, REG.RegionDescrip, CDE.CDefuncionYear
FROM PERSONA PE INNER JOIN COMUNA CO ON CO.ComunaID = PE.ComunaID
INNER JOIN PROVINCIA PRO ON PRO.ProvinciaID = CO.ProvinciaID
INNER JOIN REGION REG ON REG.RegionID = PRO.RegionID
INNER JOIN CERTIFICADO_DEFUNCION CDE ON CDE.PersonaID = PE.PersonaID
WHERE REG.RegionID NOT IN (99) AND CDE.CDefuncionYear BETWEEN 1997 AND 2017
) P PIVOT
(
COUNT (PersonaID) FOR CDefuncionYear IN ([1997],[1998], [1999],[2000],[2001],[2002],[2003],[2004],[2005],[2006],[2007],[2008],[2009],
[2010], [2011], [2012], [2013], [2014], [2015],[2016],[2017] )
) AS pvt
ORDER BY pvt.RegionDescrip;

```

Figura 22: Consulta en la base de datos de cantidad de muertes por año y región.

Tabla 52: Cantidad de muertes por año y región

REGION	AÑO DE DEFUNCION:																				
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Aisen del Gral. C. Ibáñez del Campo	370	431	403	460	422	407	423	437	441	441	445	487	529	627	574	601	493	483	498	477	525
Antofagasta	2201	2371	2319	2232	2409	2388	2368	2368	2403	2441	2799	2774	2628	2783	2686	2820	2883	3068	3033	3050	3104
Arica y Parinacota	799	868	859	870	910	922	928	950	986	963	1034	1014	1065	1135	1108	1175	1251	1205	1245	1269	1365
Atacama	1142	1107	1093	1103	1138	1163	1157	1176	1209	1261	1352	1307	1363	1394	1460	1447	1454	1484	1554	1528	1548
BioBio	7738	7773	7665	7592	7773	7726	7873	7188	8084	7763	8630	8435	8696	9256	8584	8892	9256	9440	9501	9495	9833
Coquimbo	2769	2948	2956	2952	3008	2953	3012	3257	3256	3142	3541	3419	3605	3818	3780	3986	4081	4133	4158	4297	4397
La Araucanía	5184	5171	5371	5051	5376	5336	5432	5577	5678	5581	5806	5815	5823	6231	5871	6002	6234	6148	6502	6563	6653
Libertador General Bernardo O Higgins	3994	4186	4307	4240	4350	4241	4607	4573	4702	4569	4822	4832	4720	5014	4858	5283	5162	5341	5431	5556	5440
Los Lagos	3757	3940	4234	4045	4319	4171	4280	4382	4616	4533	4731	4661	4753	4750	4691	5050	5136	5052	5108	5054	5426
Los Rios	2054	2116	2167	2116	2205	2136	2263	2311	2390	2274	2408	2371	2469	2646	2475	2633	2505	2614	2590	2690	2689
Magallanes y de La Antartica Chilena	911	853	930	928	892	936	966	937	963	954	963	1016	981	988	977	1006	960	1095	1033	1069	1087
Maule	5191	5330	5363	5228	5371	5433	5455	6579	5459	5407	5898	5834	5917	6424	5956	6333	6383	6194	6491	6573	6663
Metropolitana de Santiago	29205	30081	30812	29108	30377	30052	31158	32267	32046	32203	35317	33380	34373	36934	36396	37286	37565	38839	39084	39283	40384
Ñuble	2919	2769	2811	2753	2873	2813	2859	2962	2877	2889	3186	3023	3054	3202	3099	3195	3197	3312	3413	3308	3308
Tarapaca	889	946	946	916	991	1026	1040	1121	1092	1070	1156	1199	1196	1275	1299	1325	1448	1425	1412	1437	1384
Valparaiso	9349	9367	9748	9220	9457	9376	9851	10053	9900	10148	10912	10601	10793	11453	11171	11677	11762	12106	12274	12377	12582
<b>Total</b>	<b>71120</b>	<b>80257</b>	<b>81984</b>	<b>78814</b>	<b>81871</b>	<b>81079</b>	<b>83672</b>	<b>86138</b>	<b>86102</b>	<b>85639</b>	<b>93000</b>	<b>90168</b>	<b>91965</b>	<b>97930</b>	<b>94985</b>	<b>98711</b>	<b>99770</b>	<b>101939</b>	<b>103327</b>	<b>104026</b>	<b>106388</b>

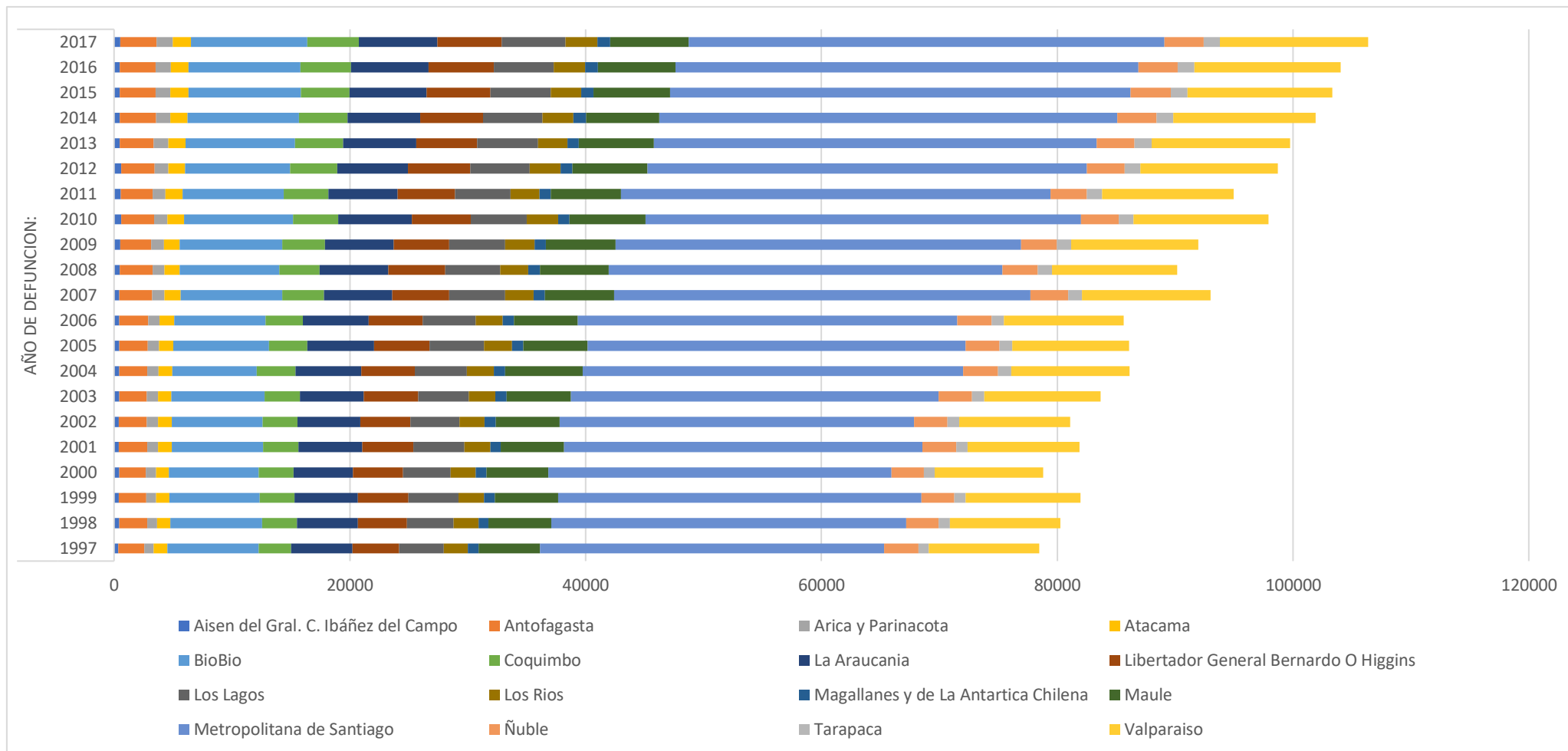


Figura 23: Cantidad de muertes por año y región.

Nota: La distribución de 1997 a 2017 se muestra por número de muertes por región.



```

SELECT REG.RegionDescrip, GRUE.CGrupoDescrip,COUNT(*) AS CANTIDAD FROM
CERTIFICADO_DEFUNCION CD INNER JOIN PERSONA P
ON CD.PersonaID=P.PersonaID INNER JOIN
COMUNA CO ON CO.ComunaID=P.ComunaID INNER JOIN
PROVINCIA PRO ON Pro.ProvinciaID=CO.ProvinciaID INNER JOIN
REGION REG ON REG.RegionID=PRO.RegionID INNER JOIN
CIE10_ENFERMEDAD EN ON CD.CEnfermedad2ID=EN.CEnfermedadID INNER JOIN
CIE10_CATEGORIA CATE ON EN.CCategoriaID=CATE.CCategoriaID INNER JOIN
CIE10_GRUPO GRUE ON GRUE.CGrupoID=CATE.CGrupoID
WHERE GRUE.CGrupoCode='X60-X84' AND CD.CDefuncionYear=1997
GROUP BY REG.RegionDescrip,GRUE.CGrupoDescrip

```

Figura 24: Consulta en la base de datos de cantidad de muertes por autolesiones y región del año 1997.

Tabla 53: Cantidad de muertes por autolesiones y región del año 1997

REGIÓN	CANTIDAD
Coquimbo	47
La Araucanía	116
Libertador General Bernardo O Higgins	66
Los Lagos	80
Los Ríos	35
Magallanes y de La Antártica Chilena	19
Maule	89
Metropolitana de Santiago	31
Ñuble	47
Tarapacá	23
Valparaíso	130
<b>Total</b>	<b>683</b>

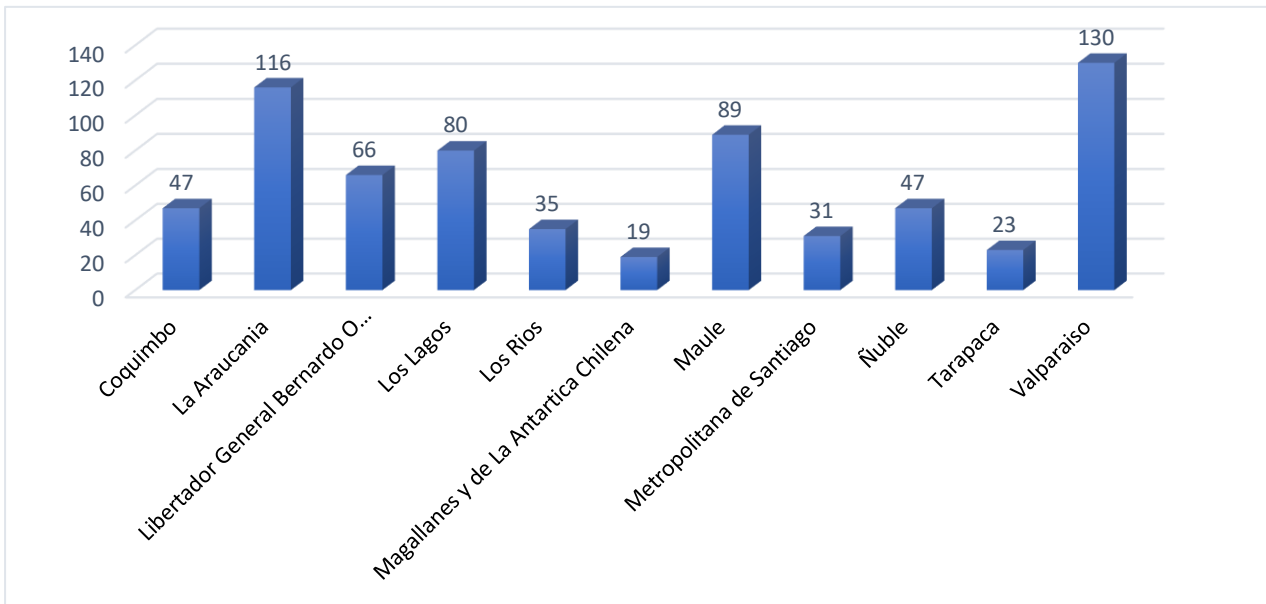


Figura 25: Cantidad de muertes por autolesiones y región del año 1997.

*Nota:* Se muestra la distribución del número de muertes por regiones ajustadas por Lesiones Autoflingidas en 1997. Donde se muestra que la tasa de mortalidad más alta se encuentra en Valparaíso.

```

SELECT C.ComunaDescrip,COUNT(*) AS 'Cantidad de muertes por comuna'
FROM PERSONA P INNER JOIN COMUNA C
ON P.ComunaID=C.ComunaID
WHERE P.Edad>9 AND P.Edad<=19 AND C.ProvinciaID=131
GROUP BY C.ComunaDescrip
HAVING COUNT(*) >200

```

Figura 26: Consulta en la base de datos de cantidad de muertes mayores a 200 de adolescentes en Santiago.

Tabla 54: Cantidad de muertes mayores a 200 de adolescentes en Santiago

Comunas	Cantidad de Defunciones	Porcentaje
Pudahuel	314	8%
Renca	265	7%
Quilicura	218	6%
La Granja	213	6%
La Pintana	406	11%
Peñalolén	306	8%
El Bosque	278	7%
Recoleta	256	7%
Maipú	603	16%
La Florida	396	11%
Santiago	270	7%
Cerro Navia	246	7%
<b>Total Comuna</b>	<b>3771</b>	<b>100 %</b>

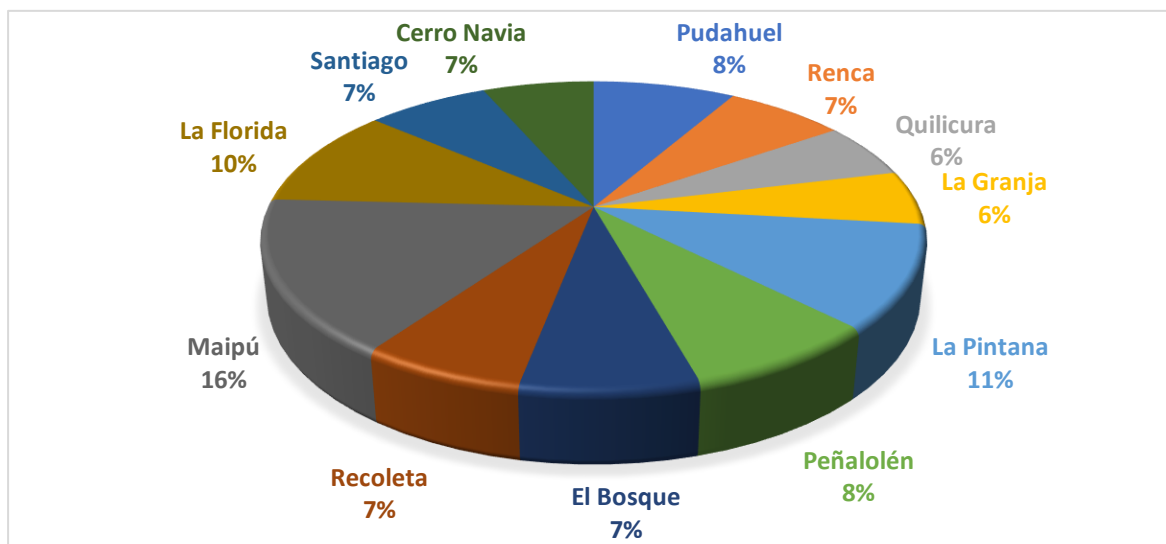


Figura 27: Cantidad de muertes mayores a 200 de adolescentes en Santiago.

Nota: La distribución de la mortalidad en adolescentes se muestra, por comuna con la provincia de Santiago y por el número de muertes mayores a 200, donde se puede ver que el porcentaje más alto (16%) de muertes corresponde a la comuna de Maipú.

```

SELECT REG.RegionDescrip, COUNT(*) AS CANTIDAD,
CONCAT (CAST (ROUND ((COUNT (*) * 100.0 / (SELECT COUNT (*)
FROM CERTIFICADO_DEFUNCION CD
)), 2) AS NUMERIC (36,2)) , '%' )
AS PORCENTAJE FROM CERTIFICADO_DEFUNCION CD INNER JOIN PERSONA P
ON CD.PersonaID=P.PersonaID INNER JOIN
COMUNA CO ON CO.ComunaID=P.ComunaID INNER JOIN
PROVINCIA PRO ON Pro.ProvinciaID=CO.ProvinciaID INNER JOIN
REGION REG ON REG.RegionID=PRO.RegionID INNER JOIN
TIPO_ZONA TP ON TP.TZonaID=P.TZonaID
WHERE (CD.CDefuncionYear BETWEEN 2000 AND 2017)
GROUP BY RegionDescrip
ORDER BY COUNT(*) DESC

```

Figura 28: Consulta en la base de datos del top 10 de la cantidad de muertes por región desde el año 2000 al 2017.

Tabla 55: Top 10 cantidad de muertes por región desde el año 2000 al 2017

Región	Cantidad	Porcentaje
Metropolitana de Santiago	626052	41%
Valparaíso	195713	13%
Biobío	154017	10%
Maule	107598	7%
La Araucanía	105679	7%
Libertador General Bernardo O Higgins	87741	6%
Los Lagos	84758	6%
Coquimbo	64795	4%
Ñuble	55323	4%
Antofagasta	48237	3%
<b>Total</b>	<b>1529913</b>	<b>100 %</b>

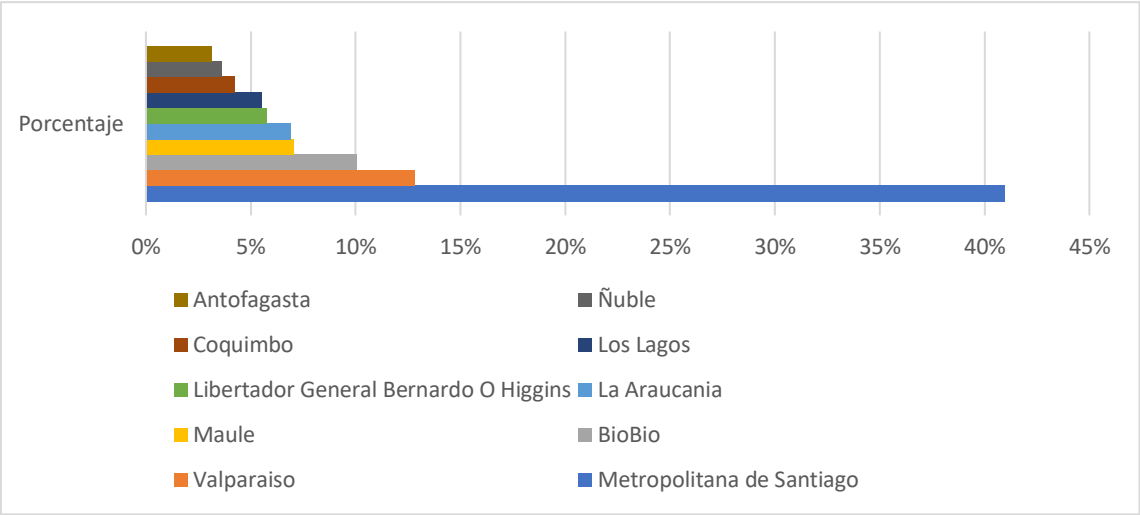


Figura 29: Cantidad de muertes por región desde el año 2000 al 2017.

Nota: Se muestra el top 10 del mayor número de muertes por región entre los rangos de los años 2000 a 2017, donde la región de Metropolitana de Santiago tiene la tasa de mortalidad más alta con un 41%.

```

SELECT TOP 5 CE.CEnfermedadDescrip AS 'Descripcion de la Enfermedad',
COUNT(*) AS 'Cantidad de muertos'
FROM CERTIFICADO_DEFUNCION CD INNER JOIN CIE10_ENFERMEDAD CE
ON CD.CEnfermedadID=CE.CEnfermedadID INNER JOIN PERSONA P
ON CD.PersonaID=P.PersonaID INNER JOIN COMUNA CO
ON CO.ComunaID=P.ComunaID INNER JOIN PROVINCIA PRO
ON Pro.ProvinciaID=CO.ProvinciaID INNER JOIN REGION REG
ON REG.RegionID=PRO.RegionID
WHERE (CD.CDefuncionYear between 2000 and 2017) and REG.RegionID=13
GROUP BY CE.CEnfermedadDescrip
ORDER BY COUNT(*) DESC

```

Figura 30: Consulta en la base de datos de las 5 causas de muerte en Santiago entre el año 2000 y 2017.

Tabla 56: Principales 5 causas de muerte en Santiago entre el año 2000 y 2017

Causa de muerte	Cantidad
Infarto agudo del miocardio, sin otra especificación.	38310
Tumor maligno de los bronquios o del pulmón, parte no especificada.	20271
Tumor maligno del estómago, parte no especificado.	18059
Accidente vascular encefálico agudo, no especificado como hemorrágico o isquémico.	17791
Neumonía, no especificada.	13827
<b>TOTAL</b>	<b>108258</b>

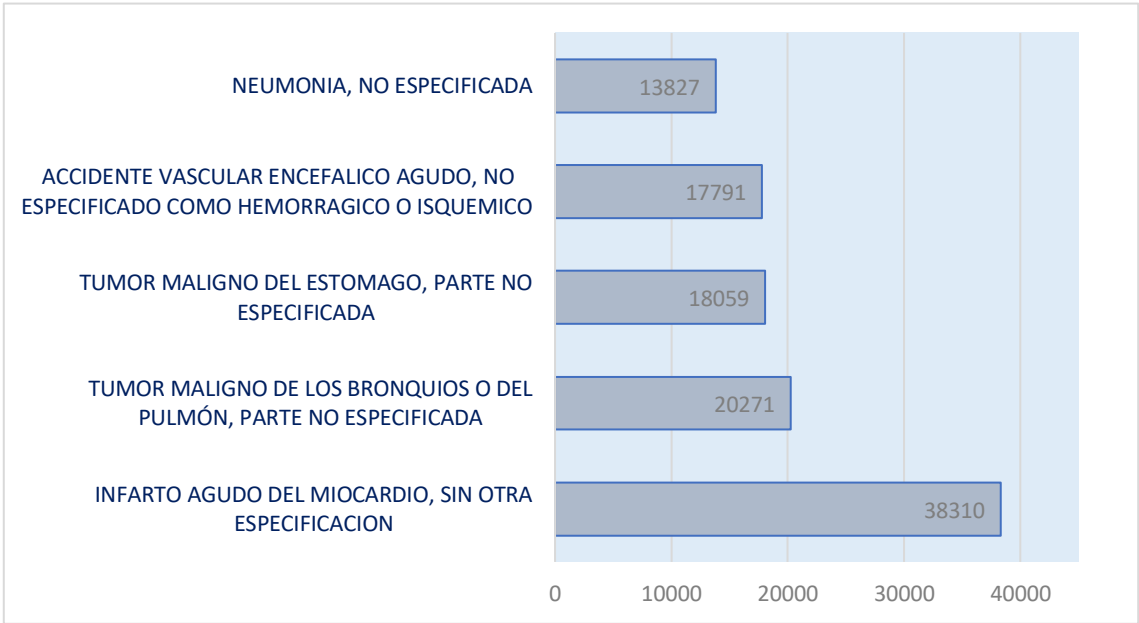


Figura 31: Las 5 causas de muerte en Santiago entre el año 2000 y 2017.

Nota: Se muestran las 5 principales enfermedades que han causado la muerte en la región metropolitana entre 2000 y 2017, donde se puede ver que la enfermedad que ha causado la mayor tasa de mortalidad es el infarto agudo.

```

SELECT COUNT(*) AS 'CANTIDAD DE MUERTE POR AÑO',
CD.CDefuncionYear FROM
PERSONA P INNER JOIN CERTIFICADO_DEFUNCION CD
ON P.PersonaID=CD.PersonaID
GROUP BY CD.CDefuncionYear
ORDER BY CD.CDefuncionYear ASC

```

Figura 32: Consulta en la base de datos de cantidad de muertes desde el año 1997 al 2017.

Tabla 57: Cantidad de muertes desde 1997 a 2017

Año	Cantidad
1997	78472
1998	80257
1999	81984
2000	78814
2001	81871
2002	81079
2003	83672
2004	86138
2005	86102
2006	85639
2007	93000
2008	90168
2009	91965
2010	97930
2011	94985
2012	98711
2013	99770
2014	101960
2015	103327
2016	104026
2017	106388
<b>Total</b>	<b>1906258</b>

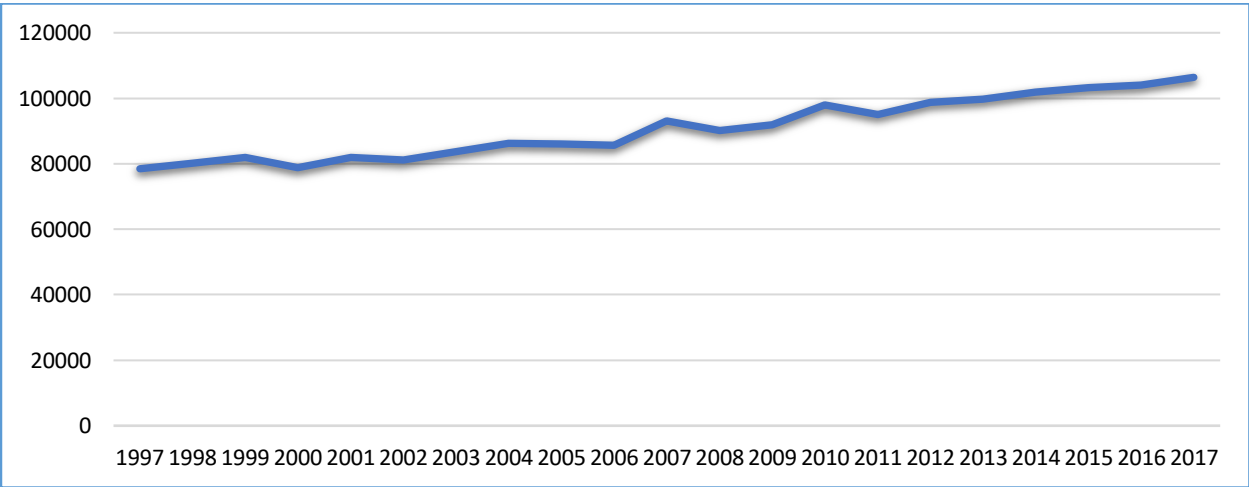


Figura 33: Cantidad de muertes desde el año 1997 al 2017.

```

SELECT COUNT(p.PersonaID) AS 'Cantidad de Personas Muertas menores de 1 año'
,EN.ENutricionalDescrip,CONCAT (CAST (ROUND ((COUNT(*) * 100.0 /
(SELECT COUNT(*) FROM PERSONA WHERE PEsMenor=1) ),2) AS NUMERIC(36,2)), '%')
AS PORCENTAJE FROM PERSONA P INNER JOIN
ESTADO_NUTRICIONAL EN ON P.ENutricionalID=EN.ENutricionalID
WHERE P.PEsMenor=1
GROUP BY EN.ENutricionalDescrip

```

Figura 34: Consulta en la base de datos de cantidad de muertes por estado nutricional de niños menores a 1 año.

Tabla 58: Cantidad de muertes por estado nutricional de niños menores de 1 año

Condicion Nutricional	Cantidad de muerte por año	Porcentaje
Desnutrición grado I	5228	12%
Desnutrición grado II	1791	4%
Desnutrición grado III	2685	6%
Eutrófico	26266	62%
Ignorado	6311	15%
<b>TOTAL</b>	<b>42281</b>	<b>100 %</b>

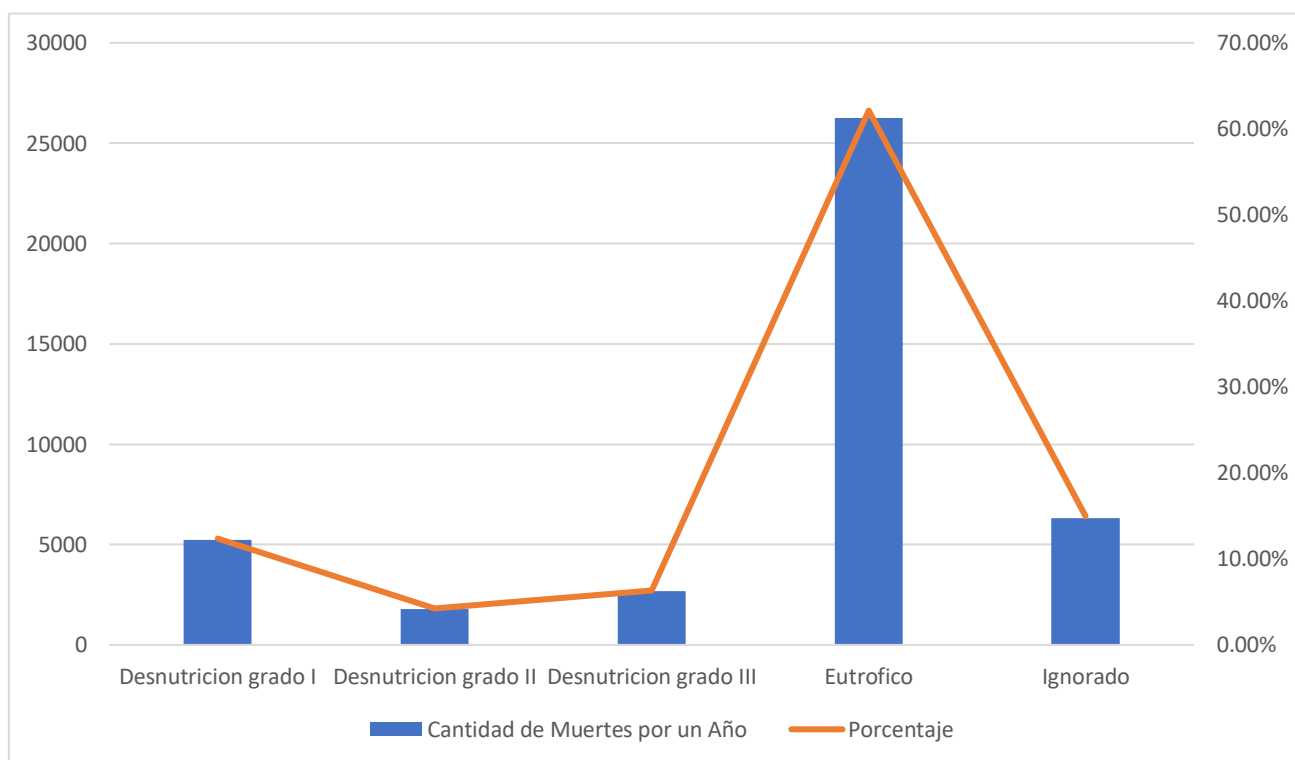


Figura 35: Cantidad de muertes por estado nutricional de niños menores a 1 año.

*Nota:* El número de muertes por estado nutricional de niños menores de 1 año se muestra como un nivel porcentual. Donde se observa que el estado predominante es eutrófico con 62%.

```

SELECT AVG(P.PEdad)AS PROMEDIO_EDAD ,COUNT(*) AS 'CANTIDAD DE MUERTOS POR CATEGORIA_OCUPACION',
CO.CocupacionalDescrip AS 'CATEGORIA_OCUPACION' FROM
PERSONA P INNER JOIN CERTIFICADO_DEFUNCION CD
ON P.PersonaID=CD.PersonaID INNER JOIN CATEGORIA_OCUPACIONAL CO
ON CO.CocupacionalID=P.CocupacionalID
WHERE(CD.CDefuncionYear BETWEEN 2000 AND 2017)
GROUP BY CO.CocupacionalDescrip

```

Figura 36: Consulta en la base de datos de las edades promedio de los fallecidos por categoría ocupacional del año 2000 al 2017.

Tabla 59: Edad promedio del fallecido por categoría ocupacional del 2000 al 2017

Categoría ocupacional	Cantidad	Edad media de defunción
Empleado	71477	47
Ignorado	27629	47
Inactivo	1405043	73
Obrero	89943	47
Patrón	3327	61
Trabajador por cuenta propia	68126	54
<b>TOTAL</b>	<b>1714137</b>	-

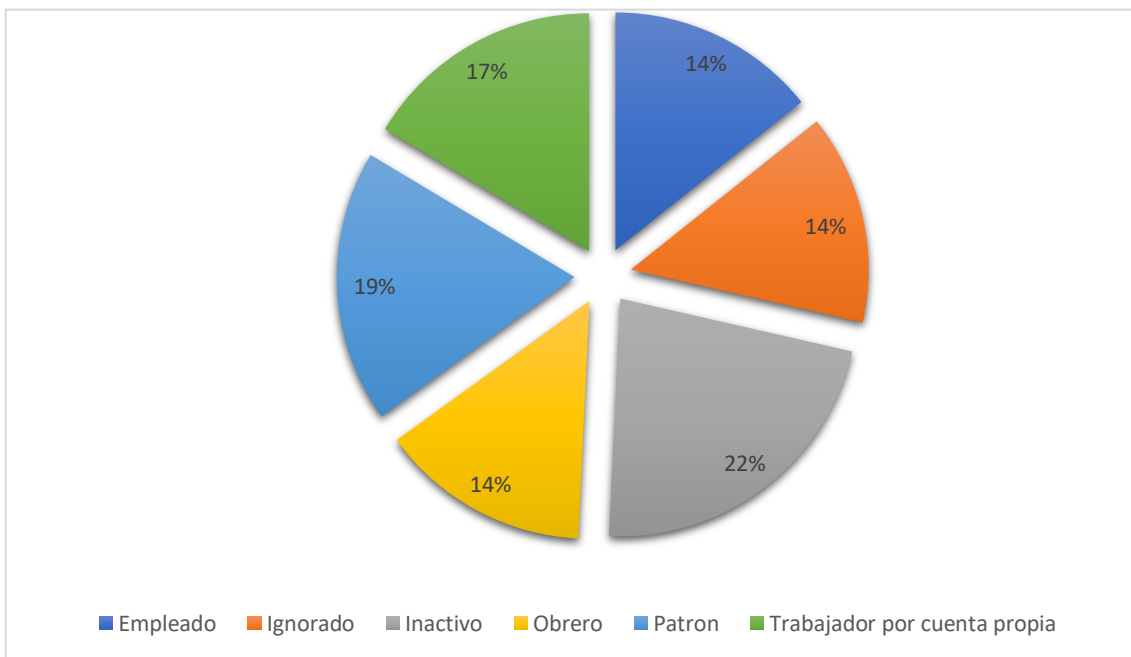


Figura 37: Edades promedio de fallecidos por categoría ocupacional.

Nota: Se muestra la edad promedio del fallecido por categoría ocupacional desde 2000 hasta 2017, donde la edad promedio de fallecimiento como mínimo es la del empleado y el trabajador.

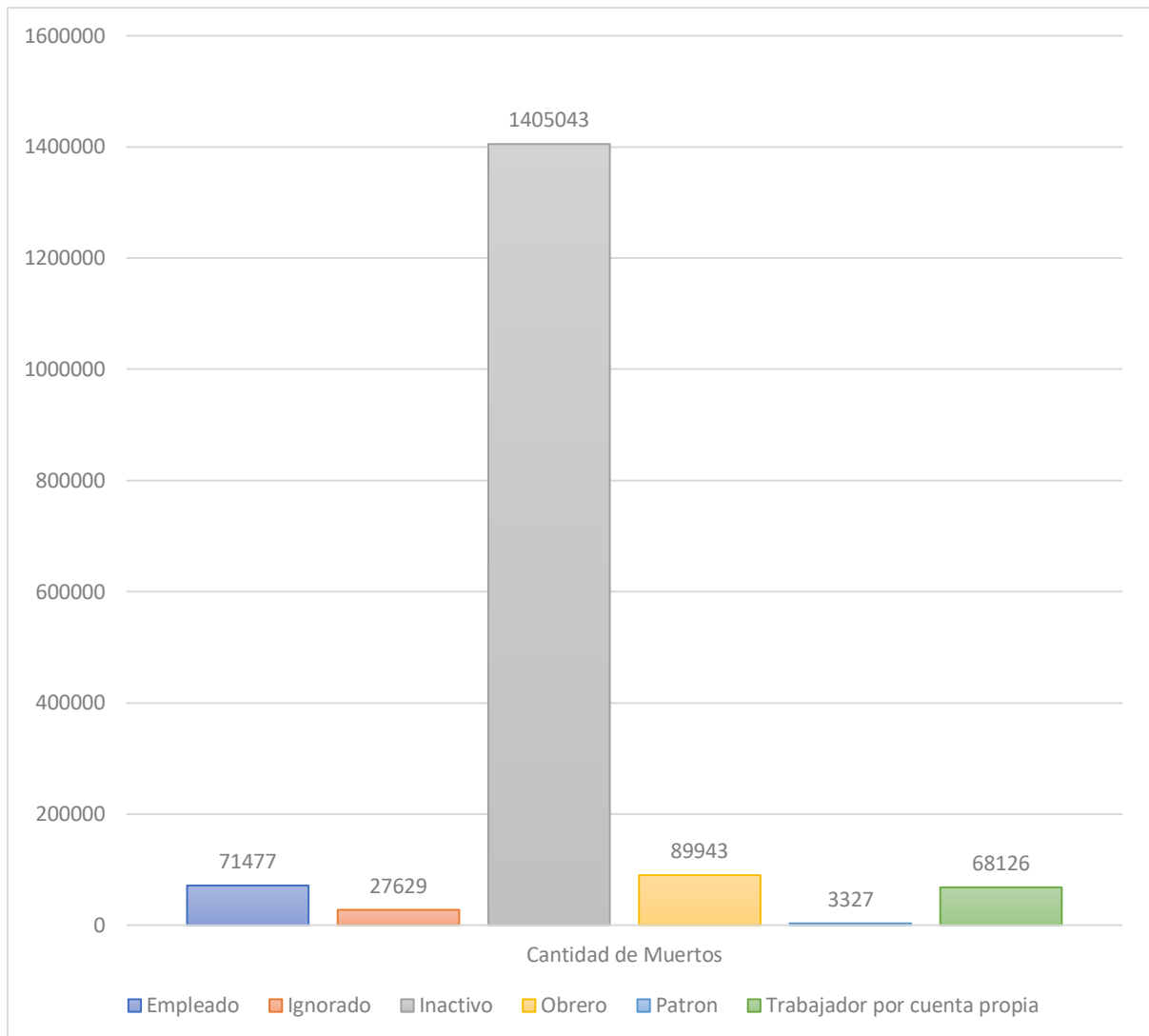


Figura 38: Edades promedio de fallecidos por categoría ocupacional.

```

SELECT AVG(P.PEdad) AS PROMEDIO_EDAD,R.RegionDescrip
FROM PERSONA P INNER JOIN COMUNA C
ON P.ComunaID=C.ComunaID INNER JOIN PROVINCIA PRO
ON PRO.ProvinciaID=C.ProvinciaID INNER JOIN REGION R
ON R.RegionID=PRO.RegionID
GROUP BY R.RegionDescrip

```

Figura 39: Consulta en la base de datos de edad promedio de los fallecidos por regiones.

*Nota:* Muestra el número de muertes por categoría ocupacional entre 2000 y 2017, donde la tasa de mortalidad más alta se encuentra en la categoría ocupacional inactiva.



Tabla 60: Edad promedio de los fallecidos por regiones

Región	Edad media
Aisén del Gral. C. Ibáñez del Campo	64
Antofagasta	65
Arica y Parinacota	67
Atacama	66
Biobío	67
Coquimbo	69
La Araucanía	69
Libertador General Bernardo O Higgins	68
Los Lagos	67
Los Ríos	68
Magallanes y de La Antártica Chilena	68
Maule	68
Metropolitana de Santiago	69
Ñuble	68
Tarapacá	65
Valparaíso	71

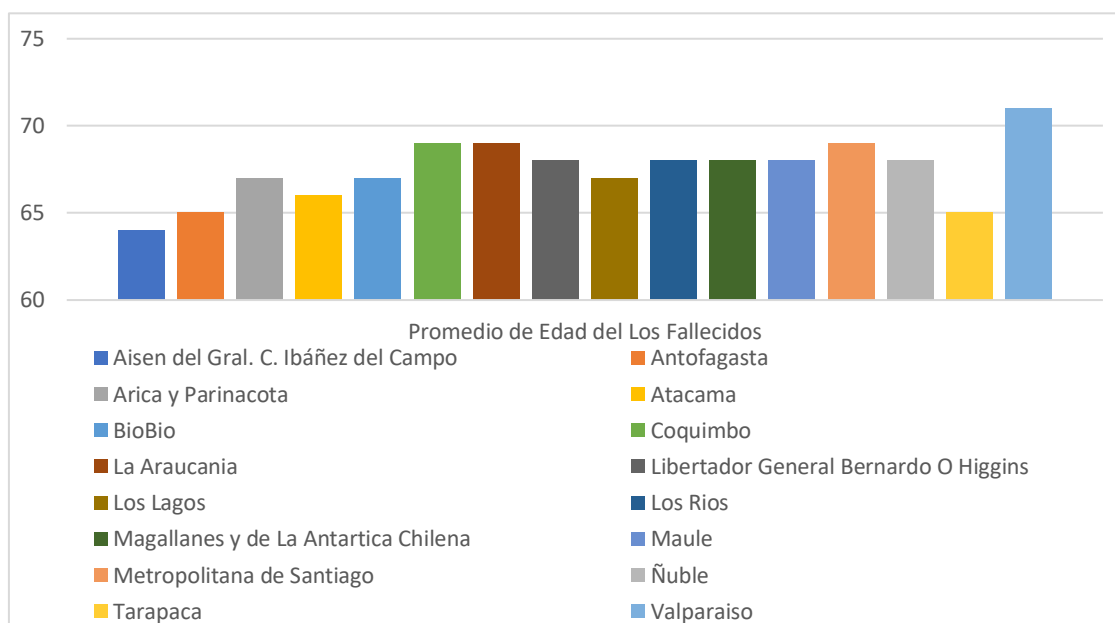


Figura 40: Edades promedio de los fallecidos por regiones.

*Nota:* Muestra el promedio de edad a la que una persona muere por región, en este caso la que tiene el promedio más alto es la región de Valparaíso con 71 y el promedio más bajo es la región de Aysén.

```

SELECT TOP 5 COUNT(*) AS CANTIDADDEMUERTOS ,
R.RegionDescrip,PRO.ProvinciaDescrip FROM
CERTIFICADO_DEFUNCION CD INNER JOIN PERSONA P
ON P.PersonaID=CD.PersonaID INNER JOIN COMUNA C
ON P.ComunaID=C.ComunaID INNER JOIN PROVINCIA PRO
ON PRO.ProvinciaID=C.ProvinciaID INNER JOIN REGION R
ON R.RegionID=PRO.RegionID
WHERE CD.CDefuncionYear=2016 AND CD.CDefuncionMes=1
GROUP BY R.RegionDescrip,PRO.ProvinciaDescrip
ORDER BY COUNT(*) ASC

```

Figura 41: Consulta en la base de datos de las 5 regiones/provincias con la mortalidad más baja en enero del 2016.

Tabla 61: Las 5 regiones/provincias con la mortalidad más baja en enero del 2016

REGIÓN	PROVINCIA	CANTIDAD
Valparaíso	Isla de Pascua	1
Aisén del Gral. C. Ibáñez del Campo	Capitán Prat	3
Aisén del Gral. C. Ibáñez del Campo	General Carrera	4
Aisén del Gral. C. Ibáñez del Campo	Aisén	8
Magallanes y de La Antártica Chilena	Tierra del Fuego	4
<b>TOTAL</b>		<b>20</b>

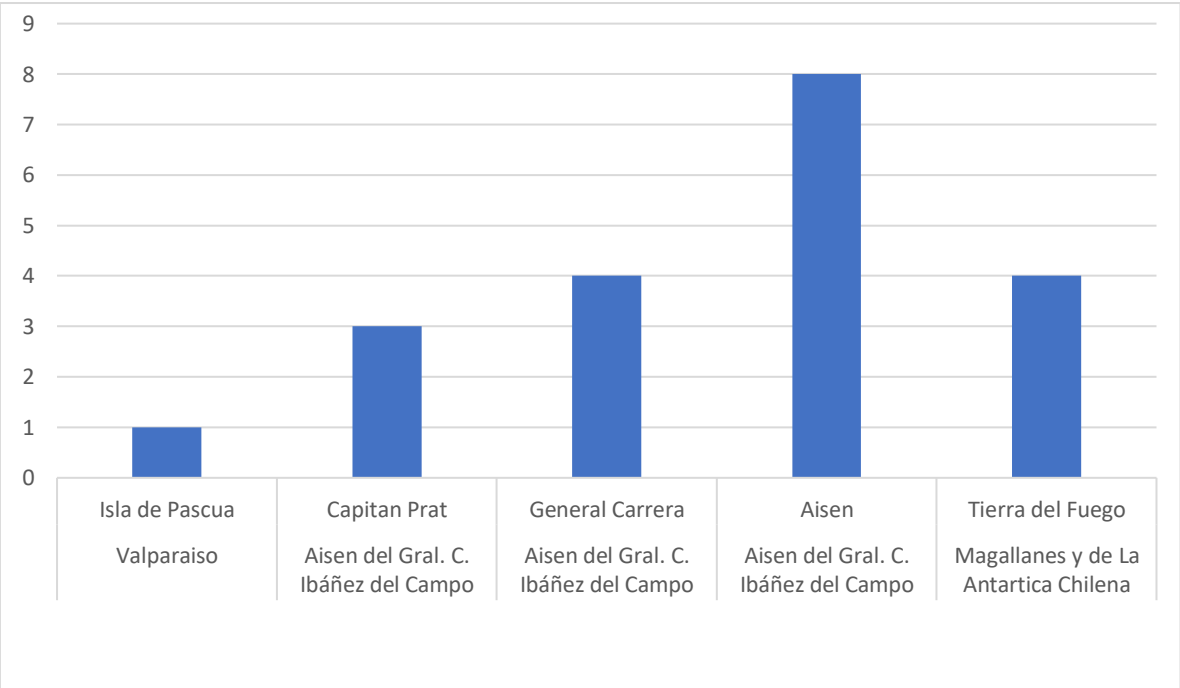


Figura 42: Las 5 regiones/provincias con la mortalidad más baja en enero 2016.

Nota: se muestran los 5 principales con la mortalidad más baja en la región y provincia en 2016 y enero, donde la tasa de mortalidad más baja en el mes de enero de 2016 tiene la provincia de Isla de Pascua en Valparaíso.

### 3. PREPARACIÓN DE DATOS

En esta fase de la metodología, los datos tienen que ser preparados para las técnicas de minería de datos que se utilizarán más adelante. La preparación implica seleccionar el subconjunto de datos que se utilizará, limpiarlo para mejorar su calidad, agregar nuevos datos a partir de los datos existentes y darle el formato solicitado por la herramienta de modelado.

#### 3.1. Seleccionar los datos

En la ejecución de esta tarea, se deben seleccionar los datos que deben considerarse para la elaboración de la estructura de datos donde se aplicará el análisis predictivo. Después de analizar los datos utilizando el resultado de las consultas SQL en las tareas de exploración y verificación de la calidad de los datos, se muestran los datos que se recomienda usar.

Tabla 62: *Preparación de los datos - Campos relevantes para el análisis*

<b>COLUMNA</b>	<b>TIPO DE DATO</b>	<b>TABLA PERTENENCIA</b>
<b>PSexo</b>	INT	PERSONA
<b>PEdad</b>	INT	PERSONA
<b>CDefuncionDia</b>	INT	CERTIFICADO_DEFUNCION
<b>CDefuncionMes</b>	INT	CERTIFICADO_DEFUNCION
<b>CDefuncionYear</b>	INT	CERTIFICADO_DEFUNCION
<b>CDefuncionFecha</b>	DATE	CERTIFICADO_DEFUNCION
<b>CEnfermedadID</b>	INT	CIE10_ENFERMEDAD
<b>CEnfermedadCode</b>	CHAR	CIE10_ENFERMEDAD
<b>CEnfermedadDescrip</b>	VARCHAR	CIE10_ENFERMEDAD
<b>CCategorialID</b>	INT	CIE10_CATEGORIA
<b>CCategoriaCode</b>	CHAR	CIE10_CATEGORIA
<b>CCategoriaDescrip</b>	VARCHAR	CIE10_CATEGORIA
<b>CGrupoID</b>	INT	CIE10_GRUPO
<b>CGrupoCode</b>	CHAR	CIE10_GRUPO

<b>COLUMNA</b>	<b>TIPO DE DATO</b>	<b>TABLA PERTENENCIA</b>
<b>CGrupoDescrip</b>	VARCHAR	CIE10_ GRUPO
<b>CCapituloID</b>	INT	CIE10_CAPITULO
<b>CCapituloCode</b>	CHAR	CIE10_CAPITULO
<b>CCapituloDescrip</b>	VARCHAR	CIE10_CAPITULO
<b>ComunalID</b>	INT	COMUNA
<b>ComunaDescrip</b>	VARCHAR	COMUNA
<b>ProvincialID</b>	INT	PROVINCIA
<b>ProvinciaDescrip</b>	VARCHAR	PROVINCIA
<b>RegiónID</b>	INT	REGIÓN
<b>RegiónDescrip</b>	VARCHAR	REGIÓN
<b>LDefuncionID</b>	INT	LOCAL_DEFUNCION
<b>LDefuncionDescrip</b>	VARCHAR	LOCAL_DEFUNCION
<b>TZonaID</b>	INT	TIPO_ZONA
<b>TZonaDescrip</b>	VARCHAR	TIPO_ZONA
<b>ECivilID</b>	INT	ESTADO_CIVIL
<b>ECivilDescrip</b>	VARCHAR	ESTADO_CIVIL
<b>TEdadID</b>	INT	TIPO_EDAD
<b>TEdadDescrip</b>	VARCHAR	TIPO_EDAD
<b>OcupacionID</b>	INT	OCUPACION
<b>OcupacionDescrip</b>	VARCHAR	OCUPACION

*Nota:* Se identificaron treinta y cuatro (34) campos relevantes para el análisis. La razón para la inclusión o exclusión de algunos campos es, como se mencionó anteriormente, la importancia de estos campos en relación con los objetivos de minería de datos que se definieron en la fase 1 (comprensión empresarial) de la metodología.

### **3.2. Limpiar los datos**

Para llevar a cabo un mejor análisis de los datos, es necesario tener los datos de una manera estructurada y limpia, por lo tanto, se procederá a eliminar los registros que contienen caracteres especiales, campos vacíos, etc. Algunos de los problemas encontrados son los siguientes:

Tabla 63: Preparación de los datos - Problemas de datos

TIPO DE PROBLEMA	DESCRIPCIÓN	SOLUCIÓN
<b>Datos ambiguos</b>	De 1990 a 1996 se descubrió que la codificación de la causa de muerte no pertenece a ninguna clasificación internacional de enfermedades, además de muchos datos faltantes	Los registros se borrarán antes de 1997, porque la causa de la muerte es una de las variables más relevantes para el análisis.
<b>Inconsistencias de codificación</b>	Duplicación de código en el campo NInstruccionID (los códigos de nivel de instrucción 0 y 5 contienen el mismo significado)	Se actualizará al código más utilizado para la descripción "ninguno", en este caso será el código 5.
<b>Datos perdidos</b>	El campo PEsMenor contiene registros vacíos, según el diccionario su valor debe ser 0 o 1 según su significado.	Cada caso se analizó por separado y se le asignó el código 0, porque este código explica que la persona es mayor de 1 año.

*Nota:* La identificación de problemas es muy importante para la limpieza de datos, ya que aclarará el panorama de los datos, esta vez se dividió en 3 tipos de problemas: datos ambiguos, inconsistencia de codificación y pérdida de datos.

A continuación, se realizarán las consultas SQL necesarias para resolver los diferentes problemas identificados en la Tabla 63.

### Datos ambiguos

Al llevar a cabo la exploración de datos, se observó que en los registros de 1990 a 1996 el campo CEnfermedadCode no contiene la codificación CIE-10:

```
SELECT DISTINCT CE.CEnfermedadCode,CE.CEnfermedadDescrip FROM CIE10_ENFERMEDAD CE
FULL OUTER JOIN CERTIFICADO_DEFUNCION CD ON CE.CEnfermedadID=CD.CEnfermedadID
WHERE CD.CDefuncionYear BETWEEN 1990 AND 1996 ORDER BY CE.CEnfermedadCode ASC
```

Figura 43: Preparación de los datos - Consulta SQL para obtener los datos ambiguos

*Nota:* El propósito de la consulta es revelar la codificación que se usó para la causa de muerte de 1990 a 1996. El resultado se puede ver en la Figura 44

	CEnfermedadCode	CEnfermedadDescrip
355	2031	NULL
356	2038	NULL
357	2039	NULL
358	2040	NULL
359	2041	NULL

	CEnfermedadCode	CEnfermedadDescrip
1	A010	FIEBRE TIFOIDEA
2	A012	FIEBRE PARATIFOIDEA B
3	A014	FIEBRE PARATIFOIDEA, NO ESPECIFICADA
4	A020	ENTERITIS DEBIDA A SALMONELLA
5	A021	SEPTICEMIA DEBIDA A SALMONELLA
6	A022	INFECCIONES LOCALIZADAS DEBIDAS A SALMONELLA

Figura 44: Preparación de datos - Resultados de datos de Ambiguos

Nota: La consulta devolvió una codificación numérica, así como una descripción nula, a diferencia de las codificaciones establecidas desde 1997, en las que, si coincide con los datos de la CIE-10, dicha codificación consta de una letra al comienzo y 3 dígitos. Por otro lado, también se observó que entre esos años los datos están incompletos, por esta razón se eligió no trabajar con los registros de esos años.

## Inconsistencias de codificación

Al realizar la consulta para verificar el número de muertes por nivel educativo, se encontró que la descripción "ninguno" pertenece a dos códigos 0 y 5. Donde el código 0 tiene un total de setenta y dos (72) registros.

```
SELECT P.NInstruccionID,N.NInstruccionDescrip, COUNT(P.NInstruccionID) AS 'Cantidad'
FROM PERSONA P INNER JOIN NIVEL_INSTRUCCION N ON P.NInstruccionID = N.NInstruccionID
GROUP BY P.NInstruccionID,N.NInstruccionDescrip ORDER BY P.NInstruccionID ASC
```

Figura 45: Preparación de los datos - Consulta SQL para obtener los códigos inconsistentes.

Nota: La consulta devuelve el número de fallecidos por nivel de instrucción educativa. El resultado se puede ver en la Figura 46.

	NInstruccionID	NInstruccionDescrip	Cantidad
1	0	Ninguno	72
2	1	Superior	90698
3	2	Medio	175313
4	3	Secundaria	354280
5	4	Basico o Primario	1027895
6	5	Ninguno	255946
7	9	Ignorado	2054

Figura 46: Preparación de los datos - Códigos inconsistentes

Nota: Los resultados muestran que el nivel de instrucción "ninguno" está duplicado por los códigos 0 y 5, donde el primero tiene 72 registros y el segundo 255946

Después de analizar, fue conveniente actualizar todos los registros encontrados con el código 0 al código 5, porque la mayoría de los registros que contienen la descripción "ninguno" contienen este código.

```
UPDATE PERSONA SET NInstruccionID = 5 FROM PERSONA WHERE NInstruccionID = 0
```

Figura 47: Preparación de los datos - Consulta SQL para actualizar los códigos

Nota: La consulta actualiza el código de nivel de instrucción "ninguno" de 0 a 5.

	NInstruccionID	NInstruccionDescrip	Cantidad
1	1	Superior	90698
2	2	Medio	175313
3	3	Secundaria	354280
4	4	Basico o Primario	1027895
5	5	Ninguno	256018
6	9	Ignorado	2054

Figura 48: Preparación de los datos - Resultados de los códigos actualizados

## Datos perdidos

El campo PEsMenor tiene dos valores: 0, que indica que el fallecido es mayor de un año y 1, lo que significa que es menor de un año. Sin embargo, se pudo observar que dentro de los registros del difunto este campo contiene varios datos vacíos.

```
SELECT PEsMenor,
(
CASE
WHEN PEsMenor = 0 THEN 'Mayor de un 1 año'
WHEN PEsMenor = 1 THEN 'Menor de un 1 año'
END
) AS 'Descripcion',COUNT(PersonaID) AS 'Cantidad' FROM PERSONA GROUP BY PEsMenor
```

Figura 49: Preparación de los datos - Consulta SQL para visualizar los datos perdidos

Nota: La consulta tiene como objetivo revelar el número de personas fallecidas mayores y menores de un año. El resultado se puede ver en la Figura 50

	PEsMenor	Descripcion	Cantidad
1			1863977
2	1	Menor de un 1 año	42281

Figura 50: Preparación de los datos - Datos perdidos

Nota: El resultado muestra que hay 42281 personas fallecidas de menos de un año, pero se muestran dos datos vacíos, donde hay una cantidad de 1863977.

La consulta devolvió el número de fallecidos en menos de un año, que corresponde al código 1, pero también devolvió datos vacíos, que corresponden a un número de 1863977 registros. Por esta razón, fue necesario analizar por qué estos datos están en blanco para resolverlos, Primero, el año de la muerte se comparó con el año de nacimiento, respaldado por el campo TEdadDescrip, que contiene el tipo de edad de las personas, ya sea en años, meses, días y horas.

```
SELECT CD.CDefuncionYear,P.YearNac,P.PEdad,TE.TEdadDescrip,P.PEsMenor FROM PERSONA P
INNER JOIN CERTIFICADO_DEFUNCION CD ON P.PersonaID =CD.PersonaID INNER JOIN TIPO_EDAD
TE ON TE.TEdadID = P.TEdadID
```

Figura 51: Preparación de los datos - Validar datos perdidos

Nota: La consulta devuelve el año de la muerte, el año de nacimiento, la edad, el tipo de edad y la condición si es menos de un año, esta consulta tiene el propósito de comparar el año de nacimiento con el año de la muerte. El resultado se puede ver en la Figura 43.

	CDefuncionYear	YearNac	PEdad	TEdadDescrip	PEsMenor
1	1999	1934	65	Edad en años	
2	1999	1920	79	Edad en años	
3	1999	1922	77	Edad en años	
4	1999	1934	65	Edad en años	
5	1998	1915	83	Edad en años	
6	1999	1910	89	Edad en años	
7	1999	1919	80	Edad en años	
8	1999	1927	72	Edad en años	
9	1999	1932	67	Edad en años	
10	1999	1916	83	Edad en años	

Figura 52: Preparación de los datos - Resultados de la validación de datos perdidos

Nota: Al realizar la consulta, se observó que el campo PEsMenor contiene datos vacíos que corresponden a los registros que, al restar el año de la muerte y el año de nacimiento, además de validarlo con la edad y el tipo de edad. Se llegó a la conclusión de que pertenecen al código 0, lo que significa que tienen más de un año.



Por las razones anteriores, elegimos actualizar los datos vacíos al código 0.

```
UPDATE PERSONA SET PEsmenor = 0 WHERE TEdadID = 1 AND PesMenor NOT IN(1)
```

*Figura 53: Preparación de los datos - Consulta SQL para actualizar los datos perdidos*

*Nota:* La consulta se actualizará al código 0 = "Mayor de un año", los registros que tienen el tipo de edad en años y no están codificados con el código 1.

### 3.3. Estructurando los datos

Para la estructuración de los datos, se ha generado una base de datos siguiendo un esquema de tipo estrella. El modelo de estrella contiene una única tabla central, que llamamos "tabla de hechos", y varias tablas de N dimensiones a su alrededor. Las tablas de dimensiones representan los factores para los cuales se analiza un área determinada del negocio, contienen los metadatos de la tabla de hechos. Por otro lado, la tabla de hechos representa el objeto del análisis, contiene los datos del estudio. El propósito de este modelo es desnormalizar las tablas. Por otro lado, este esquema es ideal por su simplicidad y velocidad cuando se usa para análisis multidimensional.

El script para crear el modelo de estrella se mostrará a continuación:

```
/* Create dimension Tabla for Certain infectious and parasitic diseases */
CREATE TABLA DIM_A00_B99 (
    ID_A00_B99 INT PRIMARY KEY CHECK (ID_A00_B99 IN (1.0)).
    Descrip_A00_B99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Tumors [Neoplasms] */
CREATE TABLA DIM_C00_D48 (
    ID_C00_D48 INT PRIMARY KEY CHECK (ID_C00_D48 IN (1.0)).
    Descrip_C00_D48 VARCHAR(10) NOT NULL
)

/* Create dimension Tabla for Diseases of the blood and hematopoietic organs. and certain disorders
that affect the Immunity mechanism */
CREATE TABLA DIM_D50_D89 (
    ID_D50_D89 INT PRIMARY KEY CHECK (ID_D50_D89 IN (1.0)).
    Descrip_D50_D89 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Endocrine. nutritional and metabolic diseases */
CREATE TABLA DIM_E00_E90 (
    ID_E00_E90 INT PRIMARY KEY CHECK (ID_E00_E90 IN (1.0)).
    Descrip_E00_E90 VARCHAR (10) NOT NULL
)
```

```

/* Create dimension Tabla for Mental and behavioral disorders */
CREATE TABLA DIM_F00_F99 (
    ID_F00_F99 INT PRIMARY KEY CHECK (ID_F00_F99 IN (1.0)).
    Descrip_F00_F99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Central nervous system diseases */
CREATE TABLA DIM_G00_G99 (
    ID_G00_G99 INT PRIMARY KEY CHECK (ID_G00_G99 IN (1.0)).
    Descrip_G00_G99 VARCHAR(10) NOT NULL
)

/* Create dimension Tabla for Diseases of the eye and its annexes */
CREATE TABLA DIM_H00_H59 (
    ID_H00_H59 INT PRIMARY KEY CHECK (ID_H00_H59 IN (1.0)).
    Descrip_H00_H59 VARCHAR(10) NOT NULL
)

/* Create dimension Tabla for Diseases of the ear and mastoid process */
CREATE TABLA DIM_H60_H95 (
    ID_H60_H95 INT PRIMARY KEY CHECK (ID_H60_H95 IN (1.0)).
    Descrip_H60_H95 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Circulatory system diseases */
CREATE TABLA DIM_I00_I99 (
    ID_I00_I99 INT PRIMARY KEY CHECK (ID_I00_I99 IN (1.0)).
    Descrip_I00_I99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Diseases of the respiratory system */
CREATE TABLA DIM_J00_J99 (
    ID_J00_J99 INT PRIMARY KEY CHECK (ID_J00_J99 IN (1.0)).
    Descrip_J00_J99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Diseases of the digestive system */
CREATE TABLA DIM_K00_K93 (
    ID_K00_K93 INT PRIMARY KEY CHECK (ID_K00_K93 IN (1.0)).
    Descrip_K00_K93 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Skin and subcutaneous tissue diseases */
CREATE TABLA DIM_L00_L99 (
    ID_L00_L99 INT PRIMARY KEY CHECK (ID_L00_L99 IN (1.0)).
    Descrip_L00_L99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Diseases of the musculoskeletal system and connective tissue*/
CREATE TABLA DIM_M00_M99 (
    ID_M00_M99 INT PRIMARY KEY CHECK (ID_M00_M99 IN (1.0)).
    Descrip_M00_M99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Diseases of the genitourinary system) */
CREATE TABLA DIM_N00_N99 (
    ID_N00_N99 INT PRIMARY KEY CHECK (ID_N00_N99 IN (1.0)).
    Descrip_N00_N99 VARCHAR(10) NOT NULL
)

```

```

/* Create dimension Tabla for Pregnancy. childbirth and the puerperium*/
CREATE TABLA DIM_O00_O99 (
    ID_O00_O99 INT PRIMARY KEY CHECK (ID_O00_O99 IN (1.0)).
    Descrip_O00_O99 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Certain conditions originating in the perinatal period*/
CREATE TABLA DIM_P00_P96 (
    ID_P00_P96 INT PRIMARY KEY CHECK (ID_P00_P96 IN (1.0)).
    Descrip_P00_P96 VARCHAR (10) NOT NULL
)

/* Create dimension Tabla for Congenital malformations. deformities. and chromosomal
abnormalities */
CREATE TABLA DIM_Q00_Q99 (
    ID_Q00_Q99 INT PRIMARY KEY CHECK (ID_Q00_Q99 IN (1.0)).
    Descrip_Q00_Q99 VARCHAR (10)
)

/* Create dimension Tabla for Symptoms. signs and abnormal clinical and laboratory findings. not
elsewhere classified */
CREATE TABLA DIM_R00_R99 (
    ID_R00_R99 INT PRIMARY KEY CHECK (ID_R00_R99 IN (1.0)).
    Descrip_R00_R99 VARCHAR (10)
)

/* Create dimension Tabla for Injuries. poisonings and some other consequences of external causes)
*/
CREATE TABLA DIM_S00_T98 (
    ID_S00_T98 INT PRIMARY KEY CHECK (ID_S00_T98 IN (1.0)).
    Descrip_S00_T98 VARCHAR (10)
)

/* Create dimension Tabla Sex
CREATE TABLA DIM_SEX (
    SexID INT PRIMARY KEY CHECK (SexID IN (1.0)).
    SexDescrip VARCHAR (25) NOT NULL
)

/* Create dimension Tabla death location*/
CREATE TABLA DEATH_LOCATION (
    DLocationID INT PRIMARY KEY CHECK (DLocationID IN (1.0)).
    DLocationDescrip VARCHAR (25) NOT NULL
)

/* Create dimension Tabla for age range from 0 to 10 */
CREATE TABLA DIM_AGE_0_10 (
    RAgeID_0_10 INT PRIMARY KEY CHECK (RAgeID_0_10 IN (1.0)).
    REdadDescrip VARCHAR (25) NOT NULL
)

/* Create dimension Tabla for age range from 11 to 19 */
CREATE TABLA DIM_AGE_11_19 (
    RAgeID_11_19 INT PRIMARY KEY CHECK (RAgeID_11_19 IN (1.0)).
    REdadDescrip VARCHAR (25) NOT NULL
)

```

/\* Create dimension Tabla for age range from 20 to 64 \*/

```
CREATE TABLA DIM_AGE_20_64 (  
    RAgeID_20_64 INT PRIMARY KEY CHECK (RAgeID_20_64 IN (1.0)).  
    REdadDescrip VARCHAR (25) NOT NULL  
)
```

/\* Create dimension Tabla for over 65 years

```
CREATE TABLA DIM_AGE_65 (  
    RAgeID_65 INT PRIMARY KEY CHECK (RAgeID_65 IN (1.0)).  
    REdadDescrip VARCHAR (25) NOT NULL  
)
```

/\* Create dimension Tabla for territorial location \*/

```
CREATE TABLA DIM_TERRITORIAL_LOCATION (  
    CommuneID INT PRIMARY KEY CHECK (CommuneID > 0).  
    CommuneDescrip VARCHAR (100).  
    PobTotalCenso2017 INT NOT NULL.  
    PobFCenso2017 INT NOT NULL.  
    PobMCenso2017 INT NOT NULL.  
    ProvinceID INT NOT NULL CHECK (ProvinceID > 0).  
    ProvinceDescrip VARCHAR (100).  
    RegiónID INT NOT NULL CHECK (RegiónID > 0).  
    RegiónDescrip VARCHAR (100) NOT NULL  
)
```

/\* Create Tabla dimension time period \*/

```
CREATE TABLA TIME_PERIOD (  
    TPeriodID INT PRIMARY KEY.  
    TYear INT NOT NULL CHECK (TYear >=1997).  
    TMonth INT NOT NULL CHECK (TMonth BETWEEN 1 AND 12).  
    TMonthName VARCHAR(50) NOT NULL.  
    TNumTrimester INT NOT NULL CHECK (TNumTrimester BETWEEN 1 AND 4).  
    TNumSemester INT NOT NULL CHECK (TNumSemester IN (1.2))  
)
```

/\* Create fact Tabla\*/

```
CREATE TABLA FACTS_DEATHS (  

```

```
    TPeriodID INT FOREIGN KEY REFERENCES TIME_PERIOD (TPeriodID).  
    CommuneID INT FOREIGN KEY REFERENCES DIM_TERRITORIAL_LOCATION  
(CommuneID).  
    DLocationID INT FOREIGN KEY REFERENCES DEATH_LOCATION (DLocationID).  
    SexID INT FOREIGN KEY REFERENCES DIM_SEX (SexID).  
    RAgeID_0_10 INT FOREIGN KEY REFERENCES DIM_AGE_0_10 (RAgeID_0_10).  
    RAgeID_11_19 INT FOREIGN KEY REFERENCES DIM_AGE_11_19 (RAgeID_11_19).  
    RAgeID_20_64 INT FOREIGN KEY REFERENCES DIM_AGE_20_64 (RAgeID_20_64).  
    RAgeID_65 INT FOREIGN KEY REFERENCES DIM_AGE_65 (RAgeID_65).  
    ID_A00_B99 INT FOREIGN KEY REFERENCES DIM_A00_B99 (ID_A00_B99).  
    ID_C00_D48 INT FOREIGN KEY REFERENCES DIM_C00_D48 (ID_C00_D48).  
    ID_D50_D89 INT FOREIGN KEY REFERENCES DIM_D50_D89 (ID_D50_D89).  
    ID_E00_E90 INT FOREIGN KEY REFERENCES DIM_E00_E90 (ID_E00_E90).  
    ID_F00_F99 INT FOREIGN KEY REFERENCES DIM_F00_F99 (ID_F00_F99).  
    ID_G00_G99 INT FOREIGN KEY REFERENCES DIM_G00_G99 (ID_G00_G99).  
    ID_H00_H59 INT FOREIGN KEY REFERENCES DIM_H00_H59 (ID_H00_H59).  
    ID_H60_H95 INT FOREIGN KEY REFERENCES DIM_H60_H95 (ID_H60_H95).  
    ID_I00_I99 INT FOREIGN KEY REFERENCES DIM_I00_I99 (ID_I00_I99).  
    ID_J00_J99 INT FOREIGN KEY REFERENCES DIM_J00_J99 (ID_J00_J99).  
    ID_K00_K93 INT FOREIGN KEY REFERENCES DIM_K00_K93 (ID_K00_K93).  
    ID_L00_L99 INT FOREIGN KEY REFERENCES DIM_L00_L99 (ID_L00_L99).
```

```
ID_M00_M99 INT FOREIGN KEY REFERENCES DIM_M00_M99 (ID_M00_M99).
ID_N00_N99 INT FOREIGN KEY REFERENCES DIM_N00_N99 (ID_N00_N99).
ID_O00_O99 INT FOREIGN KEY REFERENCES DIM_O00_O99 (ID_O00_O99).
ID_P00_P96 INT FOREIGN KEY REFERENCES DIM_P00_P96 (ID_P00_P96).
ID_Q00_Q99 INT FOREIGN KEY REFERENCES DIM_Q00_Q99 (ID_Q00_Q99).
ID_R00_R99 INT FOREIGN KEY REFERENCES DIM_R00_R99 (ID_R00_R99).
ID_S00_T98 INT FOREIGN KEY REFERENCES DIM_S00_T98 (ID_S00_T98).
Quantity INT NOT NULL
```

)

*Figura 54: Preparación de los datos - Script para la creación del modelo estrella*

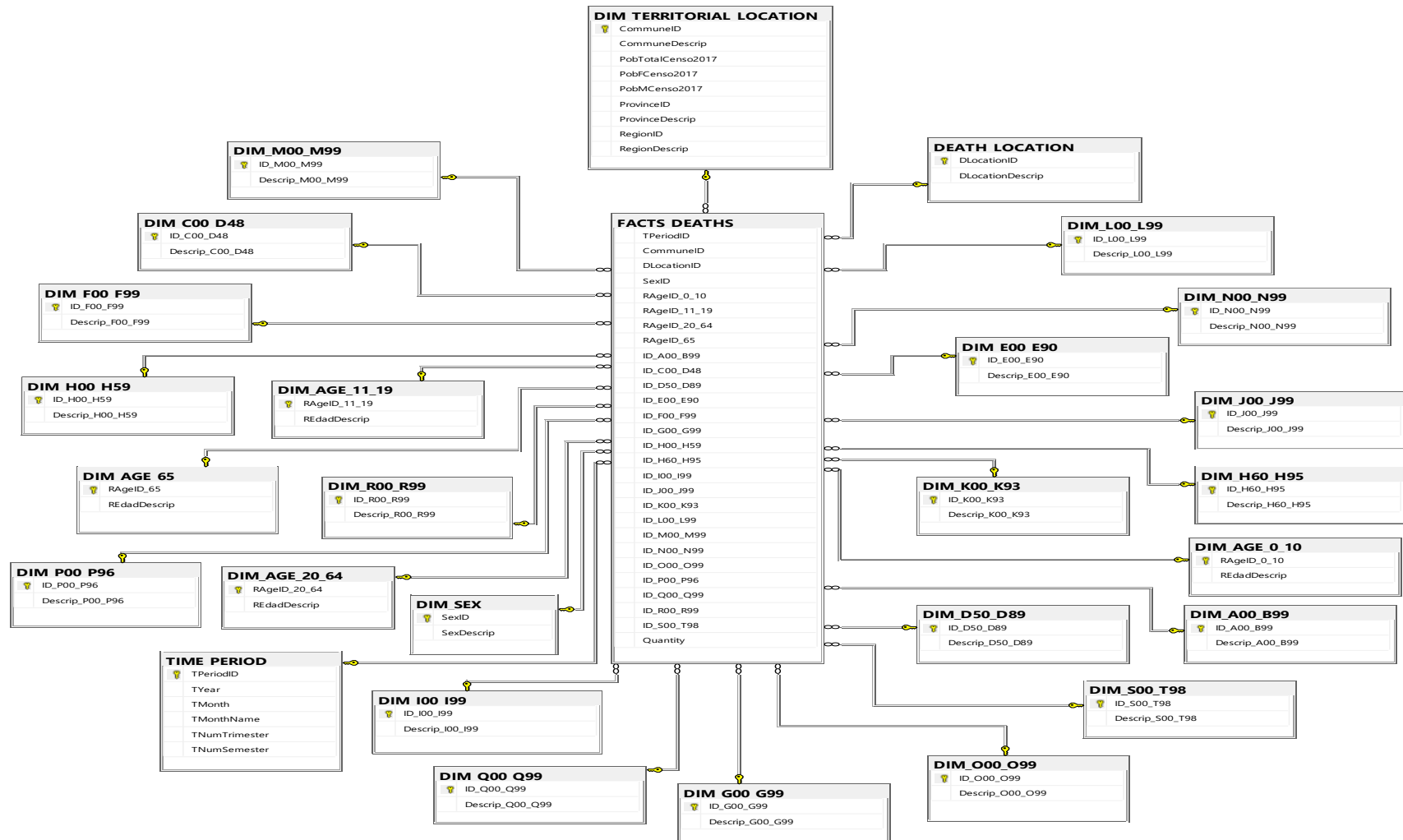


Figura 55: Preparación de los datos - Diagrama del modelo estrella.

El modelo de estrella se ha generado a partir del modelo relacional, que contiene una tabla de hechos y dimensiones, que se detallará en las siguientes tablas:

Tabla 64: *Preparación de los datos - Dimension A00\_B99*

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con ciertas enfermedades infecciosas y parasitarias con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_A00_B99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, Código: (0) Ausencia o (1) Presencia.
Descrip_A00_B99	VARCHAR(10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 65: *Preparación de los datos - Dimension C00\_D48*

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con tumores [neoplasias] con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_C00_D48	INT	Dimensión que contiene información relacionada con tumores [neoplasias] con tipos de datos ficticios.
Descrip_C00_D48	VARCHAR(10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 66: *Preparación de los datos - Dimension D50\_D89*

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con las enfermedades de la sangre y los órganos hematopoyéticos y ciertos trastornos que afectan el mecanismo de inmunidad con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_D50_D89	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_D50_D89	VARCHAR(10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 67: Preparación de los datos - Dimension E00\_E90

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con enfermedades endocrinas, nutricionales y metabólicas con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_E00_E90	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_E00_E90	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 68: Preparación de los datos - Dimension F00\_F99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con trastornos mentales y del comportamiento con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_F00_F99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_ F00_F99	VARCHAR (10)	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia

Tabla 69: Preparación de los datos - Dimension G00\_G99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con enfermedades del sistema nervioso central con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_G00_G99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_ G00_G99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.



Tabla 70: *Preparación de los datos - Dimension H00\_H59*

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con Enfermedades del ojo y sus anexos con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_H00_H59	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_H00_H59	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 71: *Preparación de los datos - Dimension H60\_H95*

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con enfermedades del oído y el proceso mastoideo con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_H60_H95	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_H60_H95	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 72: *Preparación de los datos - Dimension I00\_I99*

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con enfermedades del sistema circulatorio con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_I00_I99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_I00_I99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 73: Preparación de los datos - Dimension J00\_J99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con Enfermedades del sistema respiratorio con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_J00_J99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_J00_J99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 74: Preparación de los datos - Dimension K00\_K93

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con Enfermedades del sistema digestivo con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_K00_K93	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_K00_K93	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 75: Preparación de los datos - Dimension L00\_L99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con enfermedades de la piel y del tejido subcutáneo con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_L00_L99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_L00_L99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 76: Preparación de los datos - Dimension M00\_M99

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con enfermedades del sistema musculoesquelético y tejido conectivo con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_M00_M99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
DescripM00_M99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 77: Preparación de los datos - Dimension N00\_N99

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con Enfermedades del sistema genitourinario con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_N00_N99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_N00_N99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 78: Preparación de los datos - Dimension O00\_O99

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con el embarazo, el parto y el puerperio con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_O00_O99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_O00_O99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 79: Preparación de los datos - Dimension P00\_P96

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con ciertas condiciones originadas en el período perinatal con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_P00_P96	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
DescripP00_P96	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 80: Preparación de los datos - Dimension Q00\_Q99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con malformaciones congénitas, deformidades y anomalías cromosómicas con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_Q00_Q99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_ Q00_Q99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 81: Preparación de los datos - Dimension R00\_R99

<b>DESCRIPCIÓN</b>	Dimensión que contiene información relacionada con síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con tipos de datos ficticios.	
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_R00_R99	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
Descrip_ R00_R99	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 82: Preparación de los datos - Dimension S00\_T98

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información relacionada con lesiones, intoxicaciones y algunas otras consecuencias de causa externa con tipos de datos ficticios.		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_S00_T98	INT	Código que identifica la ausencia o presencia del capítulo de la enfermedad, código: (0) Ausencia o (1) Presencia
DescripS00_T98	VARCHAR (10)	Descripción de la ausencia o presencia del capítulo de la enfermedad.

Tabla 83: Preparación de los datos - Dimension SEXO

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información sobre el sexo biológico de la persona fallecida con tipos de datos ficticios		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
SexID	INT	Código que identifica el sexo del difunto. Si es una mujer u otra, código: (1) femenino o (0) masculino, indeterminado
SexDescrip	VARCHAR(25)	Descripción del sexo del difunto, 1 (femenino) o 0 (masculino e indeterminado)

Tabla 84: Preparación de los datos - Dimension DEATH\_LOCATION

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información sobre el lugar de la persona fallecida con tipos de datos ficticios		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
DLocationID	INT	Código que identifica el lugar de fallecimiento del fallecido. Si falleció en un hospital o clínica (1) y si fue en su hogar u otro (0)
DLocationDescrip	VARCHAR(25)	Descripción del lugar de la muerte. 1 (Hospital o Clínica) o 0 (domicilio u otro)

Tabla 85: Preparación de los datos - Dimension DIM\_AGE\_0\_10

<b>DESCRIPCIÓN</b>		
Dimensión que contiene el rango de edad de 0 a 10		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
RAgeID_0_10	INT	Código que identifica el rango de edad de 0 a 10 de los fallecidos, código: (0) Ausencia o (1) Presencia
REdadDescrip	VARCHAR (25)	Descripción del rango de muerte 0-10, si ausencia o presencia

Tabla 86: Preparación de los datos - Dimension DIM\_AGE\_11\_19

<b>DESCRIPCIÓN</b>		
Dimensión que contiene el rango de edad de 11 a 19		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
RAgeID_11_19	INT	Código que identifica el rango de edad de 11 a 19 años del fallecido, código: (0) Ausencia o (1) Presencia
REdadDescrip	VARCHAR (25)	Descripción del rango de muerte 11-19, si ausencia o presencia

Tabla 87: Preparación de los datos - Dimension DIM\_AGE\_20\_64

<b>DESCRIPCIÓN</b>		
Dimensión que contiene el rango de edad de 20 a 64		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
RAgeID_20_64	INT	Código que identifica el rango de edad de 20 a 64 años del fallecido, código: (0) Ausencia o (1) Presencia
REdadDescrip	VARCHAR (25)	Descripción del rango de muerte 20-64, si ausencia o presencia

Tabla 88: Preparación de los datos - Dimensión DIM\_AGE\_65

<b>DESCRIPCIÓN</b>		
Dimensión que contiene el rango de edad mayor de 65 años		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
RAgeID_65	INT	Código que identifica el rango de edad mayor de 65 años del fallecido, código: (0) Ausencia o (1) Presencia
REdadDescrip	VARCHAR (25)	Descripción del rango de 65 muertes, si ausencia o presencia

Tabla 89: Preparación de los datos - Dimensión TERRITORIAL\_LOCATION

<b>DESCRIPCIÓN</b>		
Dimensión que contiene información sobre el área geográfica del difunto: comuna, provincia y región		
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
CommuneID	INT	Código que identifica la comuna de residencia
CommuneDescrip	VARCHAR (100)	Descripción de la comuna
PobTotalCenso2017	INT	Cantidad de población del censo de 2017 por comuna
PobFCenso2017	INT	Cantidad de población femenina en el censo de 2017
PobMCenso2017	INT	Cantidad de población masculina en el censo de 2017
ProvinceID	INT	Código que identifica la provincia de residencia
ProvinceDescrip	VARCHAR (100)	Descripción de la provincia
RegiónID	INT	Código que identifica la región de residencia
RegiónDescrip	VARCHAR (100)	Descripción de la región.

Tabla 90: Preparación de los datos - Dimensión TIME\_PERIOD

<b>DESCRIPCIÓN</b>		Dimensión que contiene información detallada sobre la fecha de fallecimiento de la persona.
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
TPeriodID	INT	Código que identifica la fecha
TYear	INT	Número de año
TMonth	INT	Número de mes
TMonthName	VARCHAR(50)	Descripción del año
TNumTrimester	INT	Número de trimestre
TNumSemester	INT	Número de semestre

Tabla 91: Preparación de los datos - Tabla de hechos

<b>DESCRIPCIÓN</b>		La tabla contiene la información para el análisis, respaldada por las dimensiones.
<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
TPeriodID	INT	Código que identifica la hora
CommuneID	INT	Código que identifica la ubicación territorial
DLocationID	INT	Código que identifica la ubicación territorial
SexID	INT	Código que identifica el sexo biológico
RAgeID_0_10	INT	Código que identifica el rango de edad
RAgeID_11_19	INT	Código que identifica el rango de edad
RAgeID_20_64	INT	Código que identifica el rango de edad
RAgeID_20_64	INT	Código que identifica el rango de edad
ID_A00_B99	INT	Código que identifica el capítulo de la enfermedad
ID_C00_D48	INT	Código que identifica el capítulo de la enfermedad
ID_D50_D89	INT	Código que identifica el capítulo de la enfermedad



<b>DATOS</b>	<b>TIPO DE DATOS</b>	<b>DESCRIPCIÓN Y VALORES</b>
ID_E00_E90	INT	Código que identifica el capítulo de la enfermedad
ID_F00_F99	INT	Código que identifica el capítulo de la enfermedad
ID_G00_G99	INT	Código que identifica el capítulo de la enfermedad
ID_H00_H59	INT	Código que identifica el capítulo de la enfermedad
ID_H60_H95	INT	Código que identifica el capítulo de la enfermedad
ID_I00_I99	INT	Código que identifica el capítulo de la enfermedad
ID_J00_J99	INT	Código que identifica el capítulo de la enfermedad
ID_K00_K93	INT	Código que identifica el capítulo de la enfermedad
ID_L00_L99	INT	Código que identifica el capítulo de la enfermedad
ID_M00_M99	INT	Código que identifica el capítulo de la enfermedad
ID_N00_N99	INT	Código que identifica el capítulo de la enfermedad
ID_O00_O99	INT	Código que identifica el capítulo de la enfermedad
ID_P00_P96	INT	Código que identifica el capítulo de la enfermedad
ID_Q00_Q99	INT	Código que identifica el capítulo de la enfermedad
ID_R00_R99	INT	Código que identifica el capítulo de la enfermedad
ID_S00_T98	INT	Código que identifica el capítulo de la enfermedad
Quantity	INT	Número de muertes

### 3.4. Integrar los datos

Esta tarea implica cargar los datos del modelo relacional al modelo dimensional. En este sentido, los scripts creados para cargar datos para almacenar la información que se utilizará para el análisis se muestran a continuación.

```
/*LOAD DUMMY DATA TO DIMENSIONIOS*/
```

```
INSERT INTO DIM_A00_B99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_C00_D48 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_D50_D89 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_E00_E90 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_F00_F99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_G00_G99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_H00_H59 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_H60_H95 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_I00_I99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_J00_J99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_K00_K93 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_L00_L99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_M00_M99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_N00_N99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_O00_O99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_P00_P96 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_Q00_Q99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_R00_R99 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_S00_T98 VALUES (0.'Ausencia').(1.'Presencia')
```

```
INSERT INTO DIM_SEX VALUES (0.'No es femenino').(1.'Femenino')
```

```
INSERT INTO DEATH_LOCATION VALUES (0.'No es Hospital o Clinica ').(1.'Hospital o Clinica')
```

```
INSERT INTO DIM_AGE_0_10 VALUES (0.'Ausencia').(1.'Presencia')
```

```

INSERT INTO DIM_AGE_11_19 VALUES (0.'Ausencia').(1.'Presencia')

INSERT INTO DIM_AGE_20_64 VALUES (0.'Ausencia').(1.'Presencia')

INSERT INTO DIM_AGE_65 VALUES (0.'Ausencia').(1.'Presencia')

/*END OF LOADING DUMMY DATA TO DIMENSIONS*/

/*LOAD DATA TO THE TERRITORIAL LOCATION DIMENSION*/
INSERT INTO DIM_TERRITORIAL_LOCATION
(CommuneID. CommuneDescrip. PobTotalCenso2017. PobFCenso2017. PobMCenso2017.
ProvinceID. ProvinceDescrip. RegiónID. RegiónDescrip)
SELECT
    CO.ComunaID.CO.ComunaDescrip.CO.COPoblacionTotal.
    CO.COPoblacionF.CO.COPoblacionM.PRO.ProvincialID.
    PRO.ProvinciaDescrip.REG.RegiónID.REG.RegiónDescrip
FROM COMUNA CO
    INNER JOIN PROVINCIA PRO ON CO.ProvincialID = PRO.ProvincialID
    INNER JOIN REGIÓN REG ON REG.RegiónID = PRO.RegiónID
WHERE REG.RegiónID NOT IN(99)

/*LOAD DATA TO THE TIME PERIOD DIMENSION*/
INSERT INTO TIME_PERIOD
(TPeriodID.TYear.TMonth.TMonthName.TNumTrimester.TNumSemester)
SELECT DISTINCT
    YEAR(CDefuncionFecha)*100 + MONTH(CDefuncionFecha).YEAR(CDefuncionFecha).
    MONTH(CDefuncionFecha).DATENAME(MONTH. CDefuncionFecha).
    DATEPART(QUARTER. CDefuncionFecha).
    ((DATEPART(QUARTER.CDefuncionFecha)-1)/2)+1
FROM CERTIFICADO_DEFUNCION WHERE year(CDefuncionFecha) NOT IN (1900)
ORDER BY YEAR(CDefuncionFecha)*100 + MONTH(CDefuncionFecha) ASC

/*LOAD DATA TO FACTS TABLA*/
INSERT INTO FACTS_DEATHS
SELECT (CDefuncionYear*100 + CDefuncionMes) AS 'TPeriodID'.
CO.ComunaID AS 'CommuneID'.
CASE WHEN CD.LDefuncionID = 1 THEN 1 ELSE 0 END AS 'DLocationID'.
CASE WHEN PE.PSexo = 2 THEN 1 ELSE 0 END AS 'SexID'.
CASE WHEN PE.PEdad BETWEEN 0 AND 10 OR PE.TEdadID IN (2.3.4) THEN 1 ELSE 0 END AS
'RAgeID_0_10'.
CASE WHEN PE.PEdad BETWEEN 11 AND 19 AND PE.TEdadID =1 THEN 1 ELSE 0 END AS
'RAgeID_11_19' .
CASE WHEN PE.PEdad BETWEEN 20 AND 64 AND PE.TEdadID =1 THEN 1 ELSE 0 END AS
'RAgeID_20_64'.
CASE WHEN PE.PEdad >=65 AND PE.TEdadID =1 THEN 1 ELSE 0 END AS 'RAgeID_65'.
/*CAPITULOS CIE 10*/
CASE WHEN CG.CCapituloID = 1 THEN 1 ELSE 0 END AS 'A00_B99'.
CASE WHEN CG.CCapituloID = 2 THEN 1 ELSE 0 END AS 'C00_D48'.

```

```

CASE WHEN CG.CCapituloID = 3 THEN 1 ELSE 0 END AS 'D50_D89'.
CASE WHEN CG.CCapituloID = 4 THEN 1 ELSE 0 END AS 'E00_E90'.
CASE WHEN CG.CCapituloID = 5 THEN 1 ELSE 0 END AS 'F00_F99'.
CASE WHEN CG.CCapituloID = 6 THEN 1 ELSE 0 END AS 'G00_G99'.
CASE WHEN CG.CCapituloID = 7 THEN 1 ELSE 0 END AS 'H00_H59'.
CASE WHEN CG.CCapituloID = 8 THEN 1 ELSE 0 END AS 'H60_H95'.
CASE WHEN CG.CCapituloID = 9 THEN 1 ELSE 0 END AS 'I00_I99'.
CASE WHEN CG.CCapituloID = 10 THEN 1 ELSE 0 END AS 'J00_J99'.
CASE WHEN CG.CCapituloID = 11 THEN 1 ELSE 0 END AS 'K00_K93'.
CASE WHEN CG.CCapituloID = 12 THEN 1 ELSE 0 END AS 'L00_L99'.
CASE WHEN CG.CCapituloID = 13 THEN 1 ELSE 0 END AS 'M00_M99'.
CASE WHEN CG.CCapituloID = 14 THEN 1 ELSE 0 END AS 'N00_N99'.
CASE WHEN CG.CCapituloID = 15 THEN 1 ELSE 0 END AS 'O00_O99'.
CASE WHEN CG.CCapituloID = 16 THEN 1 ELSE 0 END AS 'P00_P96'.
CASE WHEN CG.CCapituloID = 17 THEN 1 ELSE 0 END AS 'Q00_Q99'.
CASE WHEN CG.CCapituloID = 18 THEN 1 ELSE 0 END AS 'R00_R99'.
CASE WHEN CG.CCapituloID = 19 THEN 1 ELSE 0 END AS 'S00_T98'.
COUNT(*) AS 'Quantity' FROM PERSONA PE
INNER JOIN CERTIFICADO_DEFUNCION CD ON PE.PersonalID = CD.PersonalID
INNER JOIN COMUNA CO ON PE.ComunalID = CO.ComunalID
INNER JOIN PROVINCIA PRO ON PRO.ProvincialID = CO.ProvincialID
INNER JOIN REGIÓN REG ON REG.RegiónID = PRO.RegiónID
INNER JOIN CIE10_ENFERMEDAD CE ON CE.CEnfermedadID = CD.CEnfermedadID
INNER JOIN CIE10_CATEGORIA CC ON CC.CCategorialID = CE.CCategorialID
INNER JOIN CIE10_GRUPO CG ON CG.CGrupID = CC.CGrupID
WHERE REG.RegiónID NOT IN (99)
AND PE.PEdad NOT IN(999)
AND CD.CDefuncionMes NOT IN(0)
GROUP BY (CDefuncionYear*100 + CDefuncionMes), CO.ComunalID, CG.CCapituloID.
CASE WHEN PE.PSexo = 2 THEN 1 ELSE 0 END.
CASE WHEN CD.LDefuncionID = 1 THEN 1 ELSE 0 END.
CASE WHEN PE.PEdad BETWEEN 0 AND 10 OR PE.TEdadID IN (2,3,4) THEN 1 ELSE 0 END.
CASE WHEN PE.PEdad BETWEEN 11 AND 19 AND PE.TEdadID =1 THEN 1 ELSE 0 END .
CASE WHEN PE.PEdad BETWEEN 20 AND 64 AND PE.TEdadID =1 THEN 1 ELSE 0 END.
CASE WHEN PE.PEdad >=65 AND PE.TEdadID =1 THEN 1 ELSE 0 END
ORDER BY (CD.CDefuncionYear*100 + CD.CDefuncionMes) ASC

```

Figura 56: Preparación de los datos – Script para cargar los datos al modelo dimensional

### 3.5. Formatear los datos

Esta tarea consiste en validar si el formato de datos es el más adecuado para la aplicación de modelos analíticos, En este caso, no es necesario llevar a cabo ningún formateo, ya que la mayoría de los datos fueron establecidos como variables dummy con 1 y 0. Tampoco es necesario cambiar el formato ni el orden de los registros, porque toda la inconsistencia de datos se resolvió en la limpieza de datos y después de cargar el modelo multidimensional.

## 4. MODELADO

Esta tarea consiste en validar si el formato de datos es el más adecuado para la aplicación de modelos analíticos. En este caso, no es necesario llevar a cabo

ningún formateo, ya que todos los datos para el análisis tienen un código de tipo entero que identifica alguna descripción o brillo. Tampoco es necesario cambiar el formato ni el orden de los registros porque toda la inconsistencia de datos se resuelve en la limpieza de datos y después de cargar el modelo multidimensional.

#### **4.1. Elegir la técnica de modelado**

Para el modelado se dispone de la herramienta Microsoft Azure Machine Learning, la que proporciona una serie de técnicas para la minería de datos, además de poder ejecutar scripts en Python y R, que son los principales lenguajes de ciencia de datos que existen en el mercado, por lo tanto, se usará algunas de las técnicas de modelado que esta herramienta ofrece de acuerdo con los objetivos de la minería de datos. De las técnicas que Microsoft Azure Machine Learning ofrece, la que mejor se adapta a los objetivos de la presente investigación sería la regresión y las técnicas de series de tiempo, ya que los problemas que queremos resolver son problemas de predicción y los campos que queremos predecir contienen valores continuos.

#### **4.2. Generar el plan de prueba**

Como paso final antes de generar los modelos, se creará un procedimiento para verificar la calidad y validez de cada técnica. Para esta verificación, tendremos en cuenta el resumen del modelo generado por la herramienta SPSS Modeler, donde se especifica una instantánea del modelo, incluida una medida de la silueta de cohesión y la distancia entre los grupos, lo que indica si los resultados son pobres, correcto o bueno. Este resumen permite verificar rápidamente si la calidad del modelo es insuficiente. Un resultado "bueno" indica que los datos resultantes de la aplicación de la técnica reflejan evidencia razonable o sólida de que existe una estructura de agrupamiento, un resultado "correcto" indica que la evidencia es débil y, en el último caso, cuando el resultado es "pobre", indica que no hay evidencia obvia.

### 4.3. Construye el modelo

En esta tarea, la herramienta de modelado debe ejecutarse en el conjunto de datos de entrenamiento. Azure Machine Learning contiene un conjunto de elementos necesarios para la construcción de un modelo, de los cuales los principales son:

- Conversiones de formato de datos
- Entrada y salida de datos
- Transformación de datos
- Selección de características
- Machine Learning
- Módulos de biblioteca OpenCV
- Módulos de lenguaje Python
- Módulos de lenguaje R

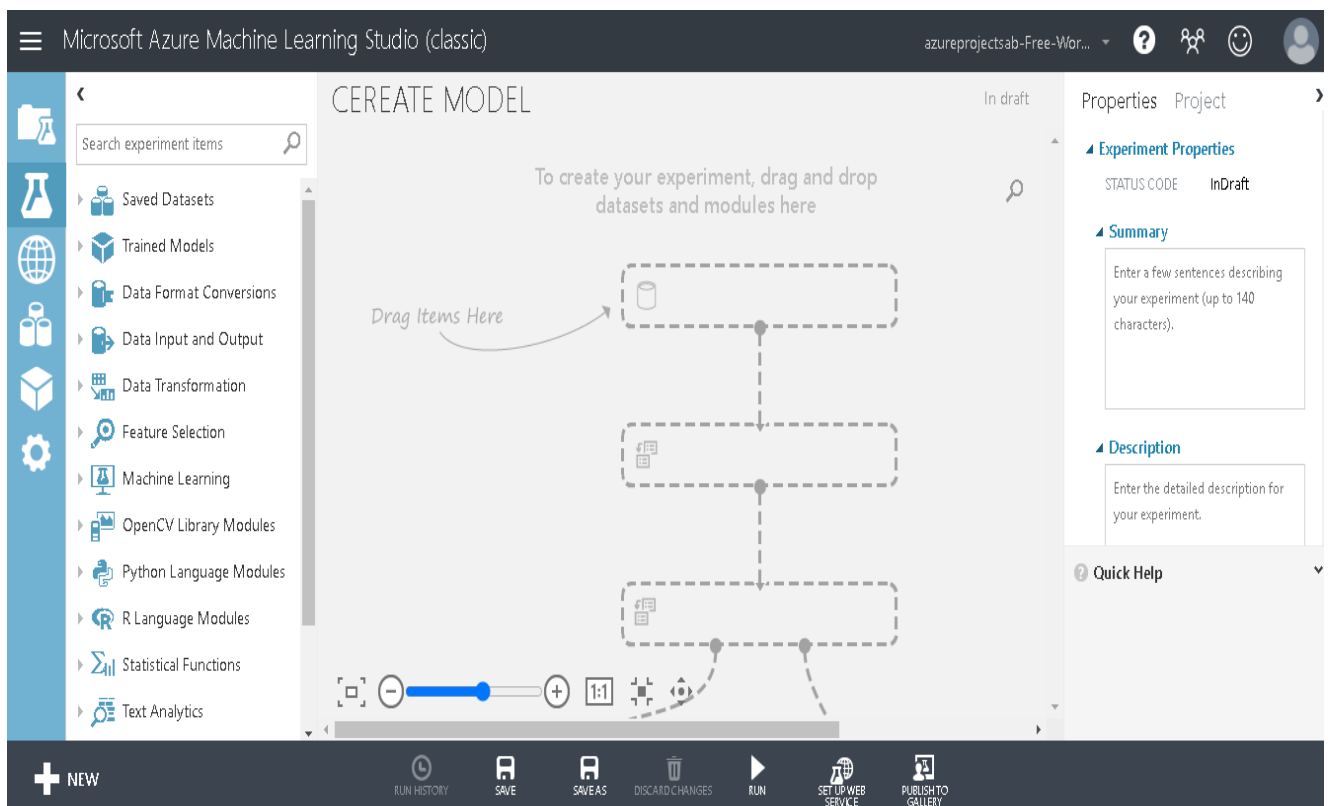


Figura 57: Modelado - Interfaz de Azure ML

Primero, los datos a modelar se cargarán desde la Base de datos SQL de Azure a través del módulo Importar datos que se encuentra en el conjunto de elementos entrada de datos y salida.

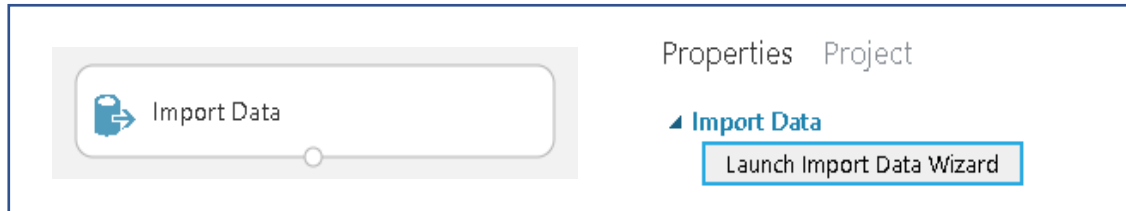


Figura 58: Modelado - Importar datos a Azure ML

*Nota:* La base de datos multidimensional se cargó en Azure SQL Database para poder usar el módulo de importación de datos y acceder a los datos rápidamente

Posteriormente, se elegirá la fuente de datos de origen. Azure Machine Learning tiene múltiples opciones de importación, pero como se explicó anteriormente, se utilizará la Base de datos SQL de Azure, por lo que los datos solicitados se completarán con el nombre del servidor, el nombre de la base de datos, nombre de usuario y contraseña, luego se realizará una prueba de conexión para validar los parámetros solicitados como se muestra en la siguiente figura

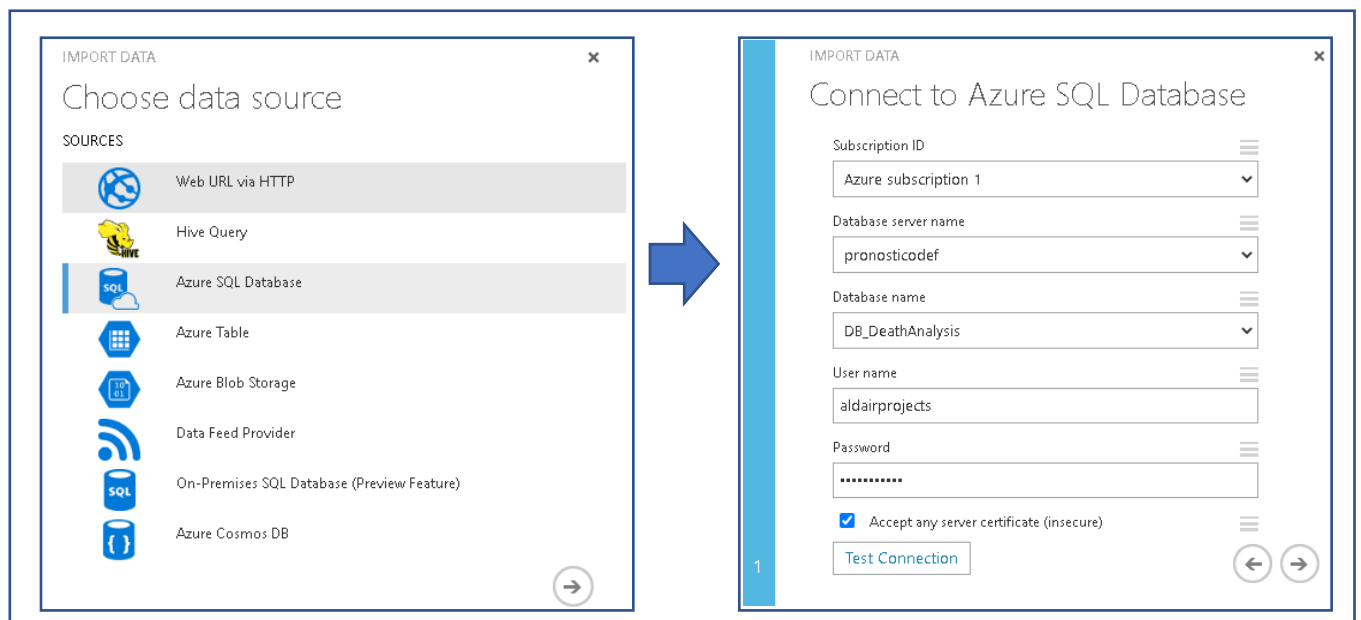


Figura 59: Modelado - Conectar la base de datos con Azure ML

*Nota:* Se puede ver que cuando se conecta a la base de datos, solicita el nombre del servidor, el nombre de la base de datos y la contraseña, luego puede realizar una prueba para validar la conexión

Para la construcción de modelos de regresión lineal múltiple, se utilizará el módulo R, que contendrá los scripts necesarios para realizar el análisis y la predicción. Las bibliotecas utilizadas para construir el modelo son:

- **Imtest:** Entre sus múltiples utilidades, en este caso se utilizará para obtener la estadística de Durbin-Watson, lo que ayudará a saber si existe una autocorrelación o no en los residuos de regresión.
- **Im.beta:** Ayudará a obtener los coeficientes estandarizados ( $\beta$ ), ya que con estos coeficientes será posible interpretar el modelo y conocer la importancia de cada variable predictiva.
- **Car:** La biblioteca Car tiene la función VIF, que mostrará el factor de inflación de la varianza para el análisis de colinealidad.

Los scripts en R utilizados para la elaboración del modelo de regresión lineal múltiple son los siguientes:

```
#=====| Map 1-based optional input ports to variables | =====
Dataset <- maml.mapInputPort(1)

# ===== | Obtaining the mean and standard deviations of each variable | =====
summary(Dataset)

SDeviation <- c( sd(Dataset$Quantity).sd(Dataset$TYear).sd(Dataset$TMonth).sd(Dataset$DLocationID).
sd(Dataset$SexID).sd(Dataset$ID_A00_B99).sd(Dataset$ID_C00_D48).sd(Dataset$ID_E00_E90).sd(Dat
aset$ID_F00_F99).sd(Dataset$ID_G00_G99).sd(Dataset$ID_I00_I99).sd(Dataset$ID_J00_J99).sd(Datas
et$ID_K00_K93).sd(Dataset$ID_L00_L99).sd(Dataset$ID_M00_M99).sd(Dataset$ID_N00_N99).sd(Datas
et$ID_O00_O99).sd(Dataset$ID_P00_P96).sd(Dataset$ID_Q00_Q99).sd(Dataset$ID_R00_R99).sd(Data
set $ID_S00_T98) )

as.data.frame( SDeviation)

#=====| Function to fit linear regression models | =====
model <- lm( Quantity ~ TYear + TMonth + DLocationID + SexID + ID_A00_B99 + ID_C00_D48 +
ID_E00_E90 + ID_F00_F99 + ID_G00_G99 + ID_I00_I99 +ID_J00_J99 + ID_K00_K93 + ID_L00_L99 +
ID_M00_M99 + ID_N00_N99 + ID_O00_O99 + ID_P00_P96 + ID_Q00_Q99 + ID_R00_R99 + ID_S00_T98.
data= Dataset)

#===== | Model Summary | =====
```



```

sum<-summary(model) # R square and Adjusted R square
dw<-dwtest(model) # Durbin -Watson
#===== | Coefficients B,  $\beta$  and VIF I =====

sum$coefficients # Non-standardized coefficients (B)
as.data.frame(lm.beta(model)$standardized.coefficients) # Standardized coefficients ( $\beta$ )
as.data.frame(vif(model)) #Variance Inflation Factor

# ===== | ANOVA I =====

anov<-summary(aov(model))
SumSQResiduals <- anov[[1]][ c("Residuals"). c("Sum Sq")] #Sum of squares of residuals
SumSQRegression <- sum(anov[[1]][ "Sum Sq"]) - SumSQResiduals #Sum of squares of regression
SumSQTotal <- SumSQResiduals + SumSQRegression #Sum of total squares
MenSQResiduals <- anov[[1]][c("Residuals"). c("Mean Sq")] #Sum of mean sq of residuals
MenSQRegression <- (sum(anov[[1]][ " Mean Sq"]) - MenSQResiduals)/20 #Sum of mean sq regression

```

Figura 60: Modelado - Script para crear el modelo de regresión lineal múltiple

Para la construcción de los modelos de regresión logística ordinal, se utilizará el módulo R, que contendrá los scripts necesarios para llevar a cabo el análisis y la predicción. Las bibliotecas utilizadas para construir el modelo son:

- **Car:** La biblioteca de Car tiene la función `anova`, que mostrará el nivel de significación de cada variable de modelo.
- **RMS:** Contiene la función `lrm`, que facilitará la aplicación de regresión logística ordinal a los datos.
- **PscI:** Ayudará a obtener los pseudo r cuadrados del modelo con el fin de mostrar la precisión del modelo.
- **MASS:** Dentro de sus múltiples utilidades, contiene una función que ofrece otra opción para aplicar la regresión logística ordinal.
- **DescTools:** Esta biblioteca contiene la función `PseudoR2` que ofrece otra opción para obtener los pseudo r cuadrados.

Se mostrarán los scripts en lenguaje R:

```

#=====| Map 1-based optional input ports to variables | =====

Dataset <- maml.mapInputPort(1)

```

```

# ===== | Obtaining the mean and standard deviations of each variable | =====

summary(Dataset) # Summary of the data
str(Dataset)
DatasetF<-within (Dataset, {Group<-as.factor(Group)}) # Convert to factor

#=====| Function to fit ordinal regression models | =====

# Method 1
model<-lrm(Group~ TYear +TMonth+ DLocationID+ SexID+ RAgeID_0_10+ RAgeID_11_19+
RAgeID_20_64+ ID_A00_B99+ ID_C00_D48+ ID_D50_D89+ ID_E00_E90+ ID_F00_F99+ ID_G00_G99+
ID_I00_I99+ ID_J00_J99+ ID_K00_K93+ ID_L00_L99+ ID_M00_M99+ ID_N00_N99+ ID_O00_O99+
ID_P00_P96+ ID_Q00_Q99+ ID_R00_R99 .data= DatasetF. tol=1e-12. maxit=20)

# Method 2
model2<-polr( Group~ TYear +TMonth+ DLocationID+ SexID+ RAgeID_0_10+ RAgeID_11_19+
RAgeID_20_64+ ID_A00_B99+ ID_C00_D48+ ID_D50_D89+ ID_E00_E90+ ID_F00_F99+ ID_G00_G99+
ID_I00_I99+ ID_J00_J99+ ID_K00_K93+ ID_L00_L99+ ID_M00_M99+ ID_N00_N99+ ID_O00_O99+
ID_P00_P96+ ID_Q00_Q99+ ID_R00_R99 .data= DatasetF. Hess=FALSE)

#===== | Model Summary | =====

summary(model2)
PseudoR2(model2. c("McFadden". "Nagel" ."CoxSnell"))
Anova(modelo2. type = 3)

```

Figura 61: Modelado - Script para crear el modelo de regresión logística ordinal

Para la construcción de los modelos de series de tiempo, se utilizará el módulo R, que contendrá los scripts necesarios para llevar a cabo el análisis y la predicción. Las bibliotecas utilizadas para construir el modelo son:

- **astsa:** Esta biblioteca acompaña el análisis de series temporales con gráficos de series y otras aplicaciones.
- **tseries:** Esta biblioteca se usará para realizar, la prueba Dickey Fuller
- **forecast:** Ayuda a pronosticar de series de tiempo o modelos de series de tiempo

```

#=====| Map 1-based optional input ports to variables | =====

Dataset <- maml.mapInputPort(1)

```

```

# ===== | Convert data to time series | =====

tsdata<-ts(Dataset.start=c(1997.1).frequency=12)
tsdata
plot(tsdata)
plot(tsdata.col="blue".main="Serie Original de Nro de Defunciones")

# ===== | Test Data | =====

adf.test(tsdata.alternative="stationary")
acf(ts(tsdata.frequency = 1)) # Correlation simple
pacf(ts(tsdata.frequency = 1)) # Correlation parcial

# ===== | Convert data to time series | =====

modelo<-auto.arima(tsdata)
modelo
summary(modelo)
coeftest(modelo)

# ===== | Evaluate model|=====

tsdiag(modelo)
Box.test(residuals(modelo). type="Ljung-Box")
#cor(fitted(modelo).tsdata)^2
forecast1<-forecast(modelo. h = 12)
forecast1
summary(forecast(forecast1))
plot(forecast1)
coeftest(forecast1)
plot(forecast1. level =95.7)

```

Figura 62: Modelado - Script para crear el modelo ARIMA

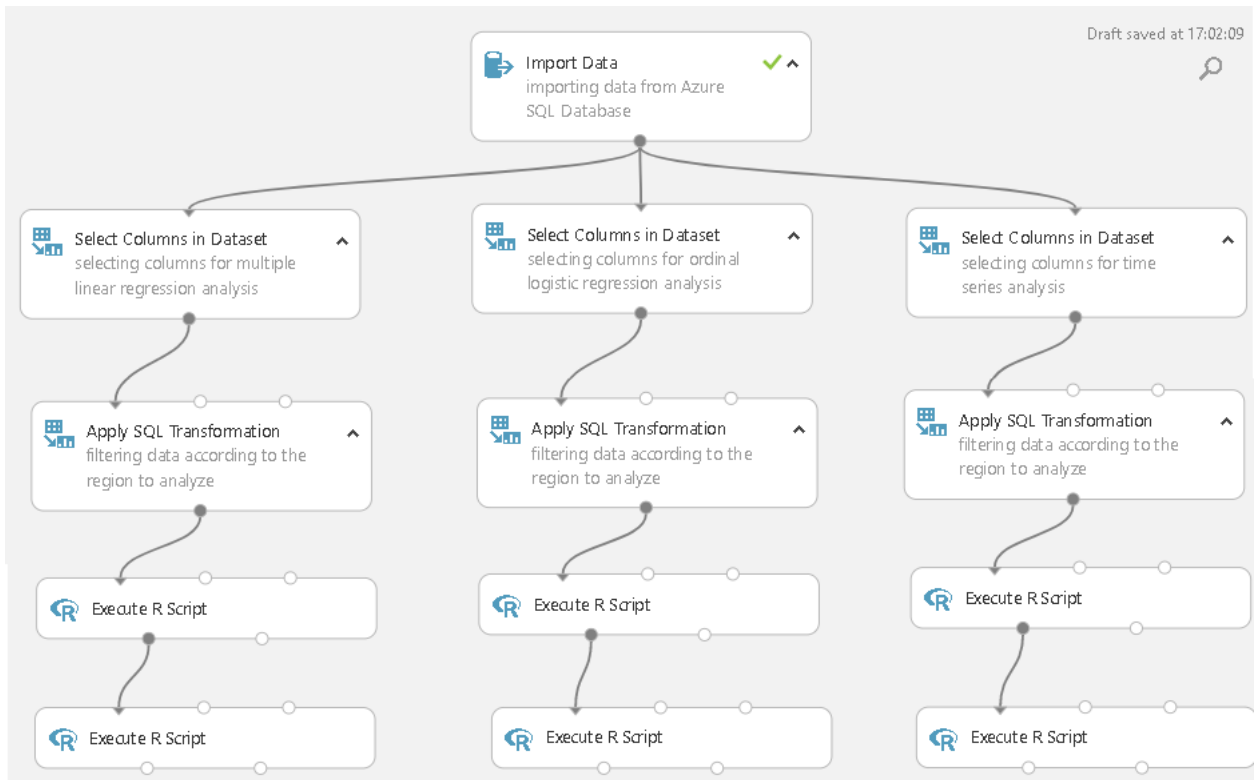


Figura 63: Modelado - Modelo en Azure ML

*Nota:* El modelo general se presenta en Azure ML. que se dividió entre las 3 técnicas que son, regresión lineal múltiple, regresión logística ordinal y series de tiempo usando los módulos de R.

## 5. EVALUACIÓN

En esta fase de la metodología se evaluarán los modelos generados, pero en esta ocasión la evaluación se hace desde el punto de vista de los objetivos de negocio en lugar de los objetivos de minería de datos. Una vez realizada esta evaluación, se debe decidir si los objetivos han sido cumplidos y de ser así se puede avanzar a la fase de implementación, de lo contrario se tendría que identificar cualquier factor que se haya podido pasar por alto y hacer una revisión del proceso.

## 5.1. Evaluación de resultados

En esta sección se mostrarán los resultados de cada técnica, debe tenerse en cuenta que el modelo se realizó para cada una de las 16 regiones de Chile, la evaluación se realizará de acuerdo con los indicadores y estadísticas mencionados en la sección 4.2 de la fase de modelado. La primera técnica utilizada para tratar de explicar la mortalidad es la regresión lineal múltiple. A continuación se mostrarán los resultados de cada región de Chile, de los cuales se mostrará la tabla resumen del modelo, la tabla ANOVA, los estadísticos descriptivos, la tabla de coeficientes y la interpretación de  $B$  y  $\beta$ .

Tabla 92: *Evaluación - Resumen del modelo de RLM de Valparaíso*

<b>Model</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Valparaíso	0.822	0.822	9.28249	1.943

Se puede ver que el modelo generado para la región de Valparaíso presenta un R cuadrado ajustado de 0.822, es decir que la varianza de la mortalidad se explica al 82.2% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados. Por lo cual, se concluye que el modelo es bueno.

Tabla 93: *Evaluación - ANOVA del modelo de RLM de Valparaíso*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	5537222.496	276861.125	3213.164	.000
Residue	6732670.752	6732670.752	-	-
Total	6732670.752	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 94: *Evaluación - Estadísticos descriptivos del modelo de Valparaíso*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	16.1329	22.01304
TYear	2007.0845	6.06591
TMonth	6.5177	3.44305
DLocationID	.5334	.49890
SexID	.5130	.49985
ID_A00_B99	.0689	.25325
ID_C00_D48	.0725	.25940
ID_E00_E90	.0725	.25940
ID_F00_F99	.0694	.25423
ID_G00_G99	.0718	.25809
ID_I00_I99	.0725	.25940
ID_J00_J99	.0725	.25940
ID_K00_K93	.0725	.25940
ID_L00_L99	.0419	.20033
ID_M00_M99	.0429	.20262
ID_N00_N99	.0723	.25904
ID_O00_O99	.0065	.08022
ID_P00_P96	.0369	.18857
ID_Q00_Q99	.0492	.21620
ID_R00_R99	.0617	.24071
ID_S00_T98	.0722	.25880

En esta tabla se puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 16.1329.

Tabla 95: *Evaluación - Coeficientes del modelo de RLM de Valparaíso*

<b>Región</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-520.885	26.075	-	-19.976	.000	-
TYear	.261	.013	.072	20.122	.000	1.001
TMonth	.110	.023	.017	4.818	.000	1.000
DLocationID	-3.057	.161	-.069	-18.975	.000	1.042
SexID	-1.446	.158	-.033	-9.127	.000	1.011
ID_A00_B99	3.064	.482	.035	6.352	.000	2.406
ID_C00_D48	53.949	.478	.636	112.929	.000	2.476
ID_E00_E90	8.070	.478	.095	16.892	.000	2.476

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_F00_F99	3.325	.482	.038	6.902	.000	2.419
ID_G00_G99	4.542	.479	.053	9.487	.000	2.461
ID_I00_I99	65.998	.478	.778	138.152	.000	2.476
ID_J00_J99	19.818	.478	.234	41.484	.000	2.476
ID_K00_K93	13.152	.478	.155	27.531	.000	2.476
ID_L00_L99	-.278	.539	-.003	-.515	.606	1.880
ID_M00_M99	.309	.536	.003	.576	.565	1.900
ID_N00_N99	5.896	.478	.069	12.335	.000	2.472
ID_O00_O99	.479	1.051	.002	.455	.649	1.147
ID_P00_P96	2.259	.559	.019	4.040	.000	1.793
ID_Q00_Q99	1.455	.518	.014	2.807	.005	2.025
ID_R00_R99	3.457	.494	.038	6.999	.000	2.279
ID_S00_T98	12.588	.478	.148	26.325	.000	2.470

Esta tabla muestra los coeficientes no estandarizados (B) que ayudarán a formar la ecuación para la predicción y los coeficientes estandarizados ( $\beta$ ), indicarán la importancia de cada variable. Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ\text{Deaths} = & - 520.88 + 0.26(\text{TYear}) + 0.11(\text{TMonth}) - 3.06 (\text{DLocationID}) - 1.45(\text{SexID}) \\ & + 3.06(\text{ID\_A00\_B99}) + 53.95(\text{ID\_C00\_D48}) + 8.07(\text{ID\_E00\_E90}) + 3.32(\text{ID\_F00\_F99}) + \\ & 4.54(\text{ID\_G00\_G99}) + 66(\text{ID\_I00\_I99}) + 19.82(\text{ID\_J00\_J99}) + 13.15(\text{ID\_K00\_K93}) - \\ & 0.28(\text{ID\_L00\_L99}) + 0.31(\text{ID\_M00\_M99}) + 5.90(\text{ID\_N00\_N99}) + 0.48(\text{ID\_O00\_O99}) + \\ & 2.26(\text{ID\_P00\_P96}) + 1.45(\text{ID\_Q00\_Q99}) + 3.46(\text{ID\_R00\_R99}) + 12.59(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B indican la relación entre el número de defunciones y cada predictor; si el valor es positivo, significa que existe una relación positiva entre el predictor y el resultado. Por otro lado, si el valor B es negativo, se concluye que

tienen una relación negativa. Sin embargo, los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.261): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.261, manteniendo las otras condiciones constantes.

**TMonth** (B = 0.110): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.110 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -3.057): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 3.057, manteniendo las otras condiciones constantes.

**SexID** (B = -1.446): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.446, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 3.064): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 3.064, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 53.949): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 53.949, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 8.070): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 8.070 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 3.325): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 3.325, manteniendo las otras condiciones constantes.



**ID\_G00\_G99** (B = 4.542): Este valor establece que, dado que hay una presencia en enfermedades del sistema nervioso central, generará un aumento de 4.542 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_I00\_I99** (B = 65.998): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 65.998, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 19.818): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 19.818 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 13.152): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 13.152 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_M00\_M99** (B = 0.309): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.309 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 5.896): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 5.896, manteniendo las otras condiciones constantes,

**ID\_O00\_O99** (B = 0.479): Este valor aclara que, dado que existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.479 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 2.259): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 2.259, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $B = 1.455$ ): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.455 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $B = 3.457$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 3.457, manteniendo constantes las otras condiciones.

**ID\_S00\_T98** ( $B = 12.588$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 12.588 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.072$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.06), la mortalidad aumenta en 7.2% desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.585 ( $0.072 * 22.013$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.017$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.44), la mortalidad aumenta en un 1.7% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.374 ( $0.017 * 22.013$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.069$ ): Este valor determina que como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 6.9% de las desviaciones estándar. La desviación estándar del número de

mueres es 22.013, por lo que esto constituye un cambio de 1.519 ( $0.069 * 22.013$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.033$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en 3.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.726 ( $0.033 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.035$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.253), la mortalidad aumenta en un 3.5% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0,770 ( $0.035 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.636$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.259), la mortalidad aumenta en un 63.6% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 14 ( $0.636 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.095$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.259), la mortalidad aumenta en un 9.5% de las desviaciones estándar, La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 2.091 ( $0.095 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.038$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.254), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.836 ( $0.038 * 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.053$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.258), la mortalidad aumenta en un 5.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.166 ( $0.053 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.778$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.259), la mortalidad aumenta en un 77.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 17.126 ( $0.778 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.234$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.259), la mortalidad aumenta en un 23.4% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 5.151 ( $0.234 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.155$ ): Este valor indica que cuando hay una presencia de enfermedades del sistema digestivo con una desviación estándar (0.259), la mortalidad aumenta en un 15.5% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 3.412 ( $0.155 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.003$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.202), la mortalidad aumenta en un 0.3% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.06 ( $0.003 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.069$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.259),

la mortalidad aumenta en un 6.9% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 1.518 ( $0.069 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.002$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.080), la mortalidad aumenta en un 0.2% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.04 ( $0.002 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.019$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.188), la mortalidad aumenta en un 1.9% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.418 ( $0.019 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.014$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.216), la mortalidad aumenta en un 1.4% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.308 ( $0.014 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.038$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.240), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 0.836 ( $0.038 \times 22.013$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.148$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.258), la mortalidad aumenta en un 14.8% de las desviaciones estándar. La desviación estándar del número de muertes es 22.013, por lo que esto constituye un cambio de 3.257 ( $0.148 \times 22.013$ ), manteniendo las otras condiciones constantes.

Tabla 96: *Evaluación - Resumen del modelo de RLM de Santiago*

Modelo	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Santiago	0.814	0.813	28.38520	1.896

Se puede ver que el modelo generado para la región de Santiago presenta un R cuadrado ajustado de 0.813, es decir que la varianza de la mortalidad se explica al 81.3% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados. Por lo cual se concluye que el modelo es bueno.

Tabla 97: *Evaluación - ANOVA del modelo de RLM de Santiago*

Modelo	Sum of squares	Square Mean	F	Sig.
Regression	53906508.684	2695325.434	3345.241	.000
Residue	12350065.848	805.719	-	-
Total	66256574.532	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 98: *Evaluación - Estadísticos descriptivos del modelo de Santiago*

Variable	Mean	Standard Deviation
Quantity	46.6546	65.70351
TYear	2007.0559	6.07455
TMonth	6.5092	3.44318
DLocationID	.5215	.49956
SexID	.5142	.49981
ID_A00_B99	.0656	.24760
ID_C00_D48	.0657	.24772
ID_E00_E90	.0657	.24772
ID_F00_F99	.0652	.24680
ID_G00_G99	.0657	.24772
ID_I00_I99	.0657	.24772
ID_J00_J99	.0657	.24772
ID_K00_K93	.0657	.24772
ID_L00_L99	.0485	.21477
ID_M00_M99	.0589	.23544
ID_N00_N99	.0657	.24772
ID_O00_O99	.0146	.11992
ID_P00_P96	.0412	.19870
ID_Q00_Q99	.0602	.23786
ID_R00_R99	.0631	.24309
ID_S00_T98	.0657	.24772

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 46.6546.

Tabla 99: *Evaluación - Coeficientes del modelo de RLM de Santiago*

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-1556.747	75.747		-20.552	.000	
TYear	.780	.038	.072	20.685	.000	1.001
TMonth	.340	.067	.018	5.102	.000	1.000
DLocationID	-14.195	.463	-.108	-30.669	.000	1.018
SexID	-2.654	.462	-.020	-5.744	.000	1.016
ID_A00_B99	12.716	1.309	.048	9.712	.000	2.002

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_C00_D48	173.912	1.309	.656	132.862	.000	2.003
ID_E00_E90	34.538	1.309	.130	26.386	.000	2.003
ID_F00_F99	15.617	1.311	.059	11.908	.000	1.995
ID_G00_G99	18.990	1.309	.072	14.507	.000	2.003
ID_I00_I99	195.076	1.309	.735	149.030	.000	2.003
ID_J00_J99	70.787	1.309	.267	54.078	.000	2.003
ID_K00_K93	50.219	1.309	.189	38.365	.000	2.003
ID_L00_L99	-.872	1.413	-.003	-.617	.537	1.755
ID_M00_M99	1.512	1.343	.005	1.125	.260	1.906
ID_N00_N99	15.879	1.309	.060	12.131	.000	2.003
ID_O00_O99	2.670	2.138	.005	1.249	.212	1.252
ID_P00_P96	11.405	1.484	.034	7.683	.000	1.657
ID_Q00_Q99	5.775	1.336	.021	4.322	.000	1.925
ID_R00_R99	11.550	1.322	.043	8.739	.000	1.966
ID_S00_T98	51.475	1.309	.194	39.324	.000	2.003

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 1556.747 + 0.78(\text{TYear}) + 0.34(\text{TMonth}) - 14.195 (\text{DLocationID}) - 2.65 \\ & (\text{SexID}) + 12.716(\text{ID\_A00\_B99}) + 173.91(\text{ID\_C00\_D48}) + 34.54(\text{ID\_E00\_E90}) + \\ & 15.61(\text{ID\_F00\_F99}) + 18.99(\text{ID\_G00\_G99}) + 195.07(\text{ID\_I00\_I99}) + 70.78(\text{ID\_J00\_J99}) \\ & + 50.219(\text{ID\_K00\_K93}) - 0.87(\text{ID\_L00\_L99}) + 1.51(\text{ID\_M00\_M99}) + 15.87 \\ & (\text{ID\_N00\_N99}) + 2.67(\text{ID\_O00\_O99}) + 11.40(\text{ID\_P00\_P96}) + 5.77(\text{ID\_Q00\_Q99}) + \\ & 11.55(\text{ID\_R00\_R99}) + 51.475(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:



**Tyear** (B = 0.780): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.780 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.340): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.340 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -14.195): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 14.195, manteniendo las otras condiciones constantes.

**SexID** (B = -2.654): Este valor establece que cuando el género es femenino, la mortalidad cae en 2.654, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 12.716): Este valor aclara que, dado que existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 12.716, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 173.912): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 173.912, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 34.538): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 34.538 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 16.617): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 16.617, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 18.990): Este valor establece que, dado que existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 18.990 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 195.076): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 195.076, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 70.787): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 70.787 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 50.219): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 50.219, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 1.512): Este valor determina que, dado que existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 1.512 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 15.879): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 15.879, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 2.670): Este valor aclara que como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 2.670 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 11.405): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 11.405 manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 5.775) Este valor indica que debido a la presencia de malformaciones congénitas, deformidades y anormalidades cromosómicas, generará un aumento de 5.775 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 11.550$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anomalías clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 11.550, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 51.475$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 51.475 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.072$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.07), la mortalidad aumenta en un 7.2% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 4.730 ( $0.072 * 65.703$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.018$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.44), la mortalidad aumenta en un 1.8% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 1.182 ( $0.018 * 65.703$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.108$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 10.8% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, entonces esto constituye un cambio de 7.095 ( $0.108 * 65.703$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.020$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 2% de las

desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 1.314 ( $0.020 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.048$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.247), la mortalidad aumenta en un 4.8% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 3.153 ( $0.048 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.656$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.247), la mortalidad aumenta en un 65.6% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 43.101 ( $0.656 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.130$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.247), la mortalidad aumenta en un 13% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 8.541 ( $0.130 * 65.703$ ), manteniendo constantes las otras condiciones.

**ID\_F00\_F99** ( $\beta = 0.059$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.246), la mortalidad aumenta en un 5.9% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 8.541 ( $0.059 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.072$ ): Este valor establece que, dado que hay enfermedades del sistema nervioso central con una desviación estándar (0.247), la mortalidad aumenta en un 7.2% de las desviaciones estándar. La desviación estándar del

número de muertes es 65.703, por lo que esto constituye un cambio de 4.730 ( $0.072 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.735$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.247), la mortalidad aumenta en un 73.5% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 48.291 ( $0.735 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.267$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.247), la mortalidad aumenta en un 26.7% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 17.542 ( $0.267 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.189$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.247), la mortalidad aumenta en un 18.9% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 12.417 ( $0.189 * 65.703$ ), manteniendo constantes las otras condiciones.

**ID\_M00\_M99** ( $\beta = 0.005$ ): Este valor determina que hay enfermedades del sistema musculo esquelético y del tejido conectivo con una desviación estándar (0.235), la mortalidad aumenta en un 0.5% de las desviaciones estándar, La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 0.328 ( $0.005 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.060$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.247), la mortalidad aumenta en un 6% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 3.942 ( $0.060 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.005$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.119), la mortalidad aumenta en un 0.5% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 0.328 ( $0.005 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.034$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.198), la mortalidad aumenta en un 3.4% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 2.233 ( $0.034 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.021$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.237), la mortalidad aumenta en un 2.1% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 1.379 ( $0.021 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.043$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.243), la mortalidad aumenta en un 4.3% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 1.379 ( $0.043 * 65.703$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.194$ ): Este valor establece que como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.247), la mortalidad aumenta en un 19.4% de las desviaciones estándar. La desviación estándar del número de muertes es 65.703, por lo que esto constituye un cambio de 12.746 ( $0.194 * 65.703$ ), manteniendo las otras condiciones constantes.

Tabla 100: *Evaluación - Resumen del modelo de RLM de Biobío*

<b>Modelo</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Biobío	0.787	0.787	7.76306	1.903

Se puede ver que el modelo generado para la región Biobío presenta un R cuadrado ajustado de 0.787, es decir que la varianza de la mortalidad se explica al 78.7% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados. En base a lo mencionado anteriormente, se concluye que el modelo es bueno.

Tabla 101: *Evaluación - ANOVA del modelo de RLM de Biobío*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig,</b>
Regression	2961127.664	148056.383	2456.752	.000
Residue	801947.615	60.265	-	-
Total	3763075.279	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo debido a que presenta un valor inferior a 0.05.

Tabla 102: *Evaluación - Estadísticos descriptivos del modelo de Biobío*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	13.2931	16.80371
TYear	2007.1261	6.05428
TMonth	6.5198	3.44506
DLocationID	.5330	.49893
SexID	.5076	.49996
ID_A00_B99	.0701	.25529
ID_C00_D48	.0756	.26442
ID_E00_E90	.0755	.26417
ID_F00_F99	.0681	.25184
ID_G00_G99	.0739	.26163

Variable	Mean	Standard Deviation
ID_I00_I99	.0756	.26442
ID_J00_J99	.0756	.26442
ID_K00_K93	.0753	.26393
ID_L00_L99	.0340	.18121
ID_M00_M99	.0386	.19274
ID_N00_N99	.0732	.26040
ID_O00_O99	.0055	.07381
ID_P00_P96	.0398	.19559
ID_Q00_Q99	.0474	.21254
ID_R00_R99	.0588	.23530
ID_S00_T98	.0753	.26381

En esta tabla se puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 13.2931

Tabla 103: *Evaluación - Coeficientes del modelo de RLM de Biobío*

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-349.781	22.328	-	-15.666	.000	-
TYear	.176	.011	.063	15.808	.000	1.002
TMonth	.107	.020	.022	5.498	.000	1.000
DLocationID	-1.520	.138	-.045	-11.019	.000	1.048
SexID	-2.534	.135	-.075	-18.754	.000	1.010
ID_A00_B99	2.806	.430	.043	6.526	.000	2.665
ID_C00_D48	43.006	.425	.677	101.293	.000	2.787
ID_E00_E90	6.152	.425	.097	14.485	.000	2.784
ID_F00_F99	2.051	.433	.031	4.741	.000	2.625
ID_G00_G99	3.489	.426	.054	8.186	.000	2.750
ID_I00_I99	45.973	.425	.723	108.279	.000	2.787
ID_J00_J99	15.327	.425	.241	36.099	.000	2.787
ID_K00_K93	13.577	.425	.213	31.958	.000	2.780
ID_L00_L99	-.106	.503	-.001	-.211	.833	1.839
ID_M00_M99	.413	.487	.005	.849	.396	1.949



Región	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_N00_N99	3.278	.427	.051	7.680	.000	2.732
ID_O00_O99	1.365	.975	.006	1.401	.161	1.145
ID_P00_P96	2.319	.485	.027	4.782	.000	1.990
ID_Q00_Q99	1.585	.465	.020	3.412	.001	2.157
ID_R00_R99	2.873	.445	.040	6.450	.000	2.428
ID_S00_T98	14.905	.425	.234	35.075	.000	2.779

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 349.781 + 0.17(\text{TYear}) + 0.10(\text{TMonth}) - 1.52 (\text{DLocationID}) - 2.53(\text{SexID}) \\ & + 2.80(\text{ID\_A00\_B99}) + 43(\text{ID\_C00\_D48}) + 6.15(\text{ID\_E00\_E90}) + 2.05(\text{ID\_F00\_F99}) + \\ & 3.48(\text{ID\_G00\_G99}) + 45.97(\text{ID\_I00\_I99}) + 15.32(\text{ID\_J00\_J99}) + 13.57(\text{ID\_K00\_K93}) - \\ & 0.10\text{ID\_L00\_L99}) + 0.41(\text{ID\_M00\_M99}) + 3.27(\text{ID\_N00\_N99}) + 1.36(\text{ID\_O00\_O99}) + \\ & 2.31(\text{ID\_P00\_P96}) + 1.58(\text{ID\_Q00\_Q99}) + 2.87(\text{ID\_R00\_R99}) + 14.90(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.176): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.176 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.107): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.107 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -1.520): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 1.520, manteniendo las otras condiciones constantes.

**SexID** (B = -2.534): Este valor establece que cuando el género es femenino, la mortalidad cae en 2.534, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 2.806): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 2.806, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 43.006): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 43.006, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 6.152): Este valor indica que, dado que existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 6.152 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 2.051): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 2.051, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 3.489): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 3.489 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 45.973): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 45.973, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 15.327): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 15.327 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 13.577): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 13.577, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.413): Este valor determina que como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.413 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 3.278): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 3.278, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 1.365): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.365 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 2.319): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 2.319, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 1.585): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.585 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 2.873): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 2.873, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 14.905): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y otras consecuencias de causas externas, generará un

aumento de 14.905 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.063$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.05), la mortalidad aumenta en un 6.3% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 1.058 ( $0.063 * 16.803$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.022$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.44), la mortalidad aumenta en un 2.2% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.369 ( $0.022 * 16.803$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.045$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en 4.5% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.756 ( $0.045 * 16.803$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.075$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae un 7.5% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 1.26 ( $0.075 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.043$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.255), la mortalidad aumenta en un 4.3% de las desviaciones estándar. La desviación

estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.722 ( $0.043 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.677$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.264), la mortalidad aumenta en un 67.7% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 11.37 ( $0.677 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.097$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.264), la mortalidad aumenta en un 9.7% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 1.629 ( $0.097 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.031$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.251), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.520 ( $0.031 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.054$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.261), la mortalidad aumenta en 5.4% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.907 ( $0.054 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.723$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.264), la mortalidad aumenta en un 72.3% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 12.148 ( $0.723 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.241$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.264), la mortalidad aumenta en un 24.1% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 4.049 ( $0.241 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.213$ ): Este valor indica que cuando hay una presencia de enfermedades del sistema digestivo con una desviación estándar (0.263), la mortalidad aumenta en un 21.3% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 3.579 ( $0.213 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.005$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.192), la mortalidad aumenta en un 0.5% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.084 ( $0.005 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.051$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.260), la mortalidad aumenta en un 5.1% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.856 ( $0.051 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.006$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.073), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.10 ( $0.006 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.027$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.073), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.10 ( $0.006 * 16.803$ ), las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.020$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.212), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.336 ( $0.020 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.040$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.235), la mortalidad aumenta en un 4% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 0.672 ( $0.040 * 16.803$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.234$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.263), la mortalidad aumenta en un 23.4% de las desviaciones estándar. La desviación estándar del número de muertes es 16.803, por lo que esto constituye un cambio de 3.931 ( $0.234 * 16.803$ ), manteniendo las otras condiciones constantes.

Tabla 104: *Evaluación - Resumen del modelo de RLM de Maule*

Modelo	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Maule	0.748	0.748	6.22050	1.944

Se puede ver que el modelo generado para la región del Maule presenta un R cuadrado ajustado de 0.748, es decir que la varianza de la mortalidad se explica al 74.8% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados. Con base en lo anterior, se concluye que el modelo es bueno.

Tabla 105: *Evaluación - ANOVA del modelo de RLM de Maule*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	1431495.842	71574.792	1849.735	.000
Residue	481051.441	38.695	-	-
Total	1912547.283	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 106: *Evaluación - Estadísticos descriptivos del modelo de Maule*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	9.9151	12.39329
TYear	2007.1235	6.05361
TMonth	6.5215	3.43340
DLocationID	.5341	.49886
SexID	.5020	.50002
ID_A00_B99	.0671	.25026
ID_C00_D48	.0809	.27276
ID_E00_E90	.0799	.27115
ID_F00_F99	.0651	.24676
ID_G00_G99	.0731	.26027
ID_I00_I99	.0809	.27276
ID_J00_J99	.0809	.27276
ID_K00_K93	.0797	.27078
ID_L00_L99	.0263	.16015
ID_M00_M99	.0323	.17675
ID_N00_N99	.0749	.26328
ID_O00_O99	.0039	.06261
ID_P00_P96	.0358	.18584
ID_Q00_Q99	.0444	.20601
ID_R00_R99	.0614	.24013
ID_S00_T98	.0804	.27189



En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es de 9.9151.

Tabla 107: *Evaluación - Coeficientes del modelo de RLM de Maule*

<b>Región</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-222.294	18.523	-	-12.001	.000	-
TYear	.112	.009	.055	12.186	.000	1.003
TMonth	.071	.016	.020	4.365	.000	1.000
DLocationID	-1.708	.115	-.069	-14.885	.000	1.054
SexID	-2.459	.112	-.099	-21.963	.000	1.009
ID_A00_B99	1.430	.376	.029	3.803	.000	2.851
ID_C00_D48	28.132	.366	.619	76.916	.000	3.203
ID_E00_E90	4.084	.366	.089	11.146	.000	3.177
ID_F00_F99	1.062	.379	.021	2.804	.005	2.811
ID_G00_G99	1.674	.371	.035	4.509	.000	3.005
ID_I00_I99	33.405	.366	.735	91.332	.000	3.203
ID_J00_J99	11.773	.366	.259	32.188	.000	3.203
ID_K00_K93	6.660	.367	.146	18.168	.000	3.170
ID_L00_L99	-.146	.462	-.002	-.315	.753	1.760
ID_M00_M99	.406	.437	.006	.929	.353	1.924
ID_N00_N99	2.020	.370	.043	5.464	.000	3.049
ID_O00_O99	1.506	.943	.008	1.598	.110	1.120
ID_P00_P96	1.420	.427	.021	3.322	.001	2.030
ID_Q00_Q99	.949	.406	.016	2.334	.020	2.256
ID_R00_R99	2.250	.383	.044	5.871	.000	2.725
ID_S00_T98	10.871	.366	.238	29.692	.000	3.189

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \mathbf{N^{\circ}Deaths} = & - 222.29 + 0.11(\mathbf{TYear}) + 0.07(\mathbf{TMonth}) - 1.70 (\mathbf{DLocationID}) - 2.45(\mathbf{SexID}) \\ & + 1.43(\mathbf{ID\_A00\_B99}) + 28.13(\mathbf{ID\_C00\_D48}) + 4.08(\mathbf{ID\_E00\_E90}) + 1.06(\mathbf{ID\_F00\_F99}) + \\ & 1.67(\mathbf{ID\_G00\_G99}) + 33.4(\mathbf{ID\_I00\_I99}) + 11.77(\mathbf{ID\_J00\_J99}) + 6.66(\mathbf{ID\_K00\_K93}) - \\ & 0.14(\mathbf{ID\_L00\_L99}) + 0.40(\mathbf{ID\_M00\_M99}) + 2.02(\mathbf{ID\_N00\_N99}) + 1.50(\mathbf{ID\_O00\_O99}) + \\ & 1.42(\mathbf{ID\_P00\_P96}) + 0.94(\mathbf{ID\_Q00\_Q99}) + 2.25(\mathbf{ID\_R00\_R99}) + 10.87(\mathbf{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.112): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.112 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.071): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de la mortalidad de 0.071, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 1.708): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 1.708, manteniendo las otras condiciones constantes.

**SexID** (B = - 2.459): Este valor establece que cuando el género es femenino, la mortalidad cae en 2.459, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 1.430): Este valor aclara que, dado que hay presencia de ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 1.430, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 28.132): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 28.132, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 4.084): Este valor indica que, dado que existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 4.084 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 1.062): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 1.062, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 1.674): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 1.674 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 33.405): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 33.405, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 11.773): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 11.773, problemas de otras condiciones constantes.

**ID\_K00\_K93** (B = 6.660): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 6.660, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.406): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.406 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 2.020): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 2.020, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 1.506): Este valor aclara que como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.506 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 1.420): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 1.420, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.949): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.949 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_R00\_R99** (B = 2.250): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificadas en otra parte, la mortalidad permanece en 2.250, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 10.871): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 10.871 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.055$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.05), la mortalidad aumenta en 5.5% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.681 ( $0.055 * 12.393$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.020$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.43), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.247 ( $0.020 * 12.393$ ), las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.069$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 6.9% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.855 ( $0.069 * 12.393$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.099$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 9.9% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 1.226 ( $0.099 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.029$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.250), la mortalidad aumenta en un 2.9% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.359 ( $0.029 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.619$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.272), la mortalidad aumenta en un 61.9% de desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 7.67 ( $0.619 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.089$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.271), la mortalidad aumenta en un 8.9% de las desviaciones

estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 1.102 ( $0.089 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.021$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.246), la mortalidad aumenta en un 2.1% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.260 ( $0.021 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.035$ ): Este valor establece que, dado que hay enfermedades del sistema nervioso central con una desviación estándar (0.260), la mortalidad aumenta en un 3.5% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.433 ( $0.035 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.735$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.272), la mortalidad aumenta en un 73.5% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 9.108 ( $0.735 * 12.393$ ), manteniendo constantes las otras condiciones.

**ID\_J00\_J99** ( $\beta = 0.259$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.272), la mortalidad aumenta en un 25.9% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 3.209 ( $0.259 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.146$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.270), la mortalidad aumenta en un 14.6% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 1.809 ( $0.146 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.006$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.176), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.074 ( $0.006 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.043$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.263), la mortalidad aumenta en un 4.3% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.532 ( $0.043 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.008$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.062), la mortalidad aumenta en un 0.8% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.10 ( $0.008 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.021$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.185), la mortalidad aumenta en un 2.1% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.26 ( $0.021 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.016$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.206), la mortalidad aumenta en un 1.6% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.198 ( $0.016 * 12.393$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.044$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.240), la mortalidad aumenta en un 4.4% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 0.545 ( $0.044 * 12.393$ ), manteniendo constantes las demás condiciones.

**ID\_S00\_T98** ( $\beta = 0.238$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.271), la mortalidad aumenta en un 23.8% de las desviaciones estándar. La desviación estándar del número de muertes es 12.393, por lo que esto constituye un cambio de 2.949 ( $0.238 * 12.393$ ), manteniendo las otras condiciones constantes.

Tabla 108: *Evaluación - Resumen del modelo de RLM de O'Higgins*

Modelo	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
O'Higgins	0.722	0.721	5.05957	1.943

Se puede ver que el modelo generado para la región del Libertador General Bernardo O'Higgins presenta una R cuadrado ajustado de 0.721, es decir que la varianza de la mortalidad se explica al 72.1% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados. En base a lo mencionado anteriormente, se concluye que el modelo es bueno.

Tabla 109: *Evaluación - ANOVA del modelo de RLM de O'Higgins*

Modelo	Sum of squares	Square Mean	F	Sig.
Regression	802896.328	40144.816	1568.204	.000
Residue	309801.884	25.599	-	-
Total	1112698.212	-	-	-



La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 110: *Evaluación - Estadísticos descriptivos del modelo de O'Higgins*

Variable	Mean	Standard Deviation
Quantity	8.2669	9.58080
TYear	2007.1410	6.04699
TMonth	6.5292	3.43029
DLocationID	.5408	.49835
SexID	.5004	.50002
ID_A00_B99	.0642	.24508
ID_C00_D48	.0831	.27612
ID_E00_E90	.0814	.27348
ID_F00_F99	.0648	.24625
ID_G00_G99	.0714	.25756
ID_I00_I99	.0831	.27612
ID_J00_J99	.0831	.27612
ID_K00_K93	.0812	.27310
ID_L00_L99	.0273	.16297
ID_M00_M99	.0291	.16814
ID_N00_N99	.0722	.25879
ID_O00_O99	.0056	.07469
ID_P00_P96	.0351	.18393
ID_Q00_Q99	.0415	.19943
ID_R00_R99	.0652	.24697
ID_S00_T98	.0813	.27323

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 8.266

Tabla 111: *Evaluación - Coeficientes del modelo de RLM de O'Higgins*

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-228.999	15.282	-	-14.985	.000	-
TYear	.116	.008	.073	15.189	.000	1.003
TMonth	.055	.013	.020	4.106	.000	1.001
DLocationID	-1.571	.095	-.082	-16.514	.000	1.065

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
SexID	-2.003	.092	-.105	-21.675	.000	1.011
ID_A00_B99	1.101	.321	.028	3.435	.001	2.922
ID_C00_D48	21.249	.309	.612	68.780	.000	3.446
ID_E00_E90	3.231	.310	.092	10.430	.000	3.398
ID_F00_F99	.865	.321	.022	2.698	.007	2.953
ID_G00_G99	1.359	.316	.037	4.304	.000	3.129
ID_I00_I99	25.122	.309	.724	81.317	.000	3.446
ID_J00_J99	9.920	.309	.286	32.108	.000	3.446
ID_K00_K93	6.253	.310	.178	20.179	.000	3.392
ID_L00_L99	-.105	.384	-.002	-.274	.784	1.851
ID_M00_M99	.163	.377	.003	.432	.666	1.907
ID_N00_N99	1.787	.315	.048	5.675	.000	3.144
ID_O00_O99	1.182	.670	.009	1.764	.078	1.185
ID_P00_P96	1.387	.362	.027	3.833	.000	2.097
ID_Q00_Q99	.858	.348	.018	2.467	.014	2.276
ID_R00_R99	2.639	.321	.068	8.223	.000	2.975
ID_S00_T98	8.401	.310	.240	27.104	.000	3.396

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 228.999 + 0.11(\text{TYear}) + 0.05(\text{TMonth}) - 1.57(\text{DLocationID}) - 2(\text{SexID}) + \\ & 1.10(\text{ID\_A00\_B99}) + 21.24(\text{ID\_C00\_D48}) + 3.23(\text{ID\_E00\_E90}) + 0.86(\text{ID\_F00\_F99}) + \\ & 1.35(\text{ID\_G00\_G99}) + 25.12(\text{ID\_I00\_I99}) + 9.92(\text{ID\_J00\_J99}) + 6.25(\text{ID\_K00\_K93}) - \\ & 0.10(\text{ID\_L00\_L99}) + 0.16(\text{ID\_M00\_M99}) + 1.78(\text{ID\_N00\_N99}) + 1.18(\text{ID\_O00\_O99}) + \\ & 1.38(\text{ID\_P00\_P96}) + 0.85(\text{ID\_Q00\_Q99}) + 2.63(\text{ID\_R00\_R99}) + 8.40(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.116): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.116 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.055): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.05 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 1.571): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 1.571, manteniendo las otras condiciones constantes.

**SexID** (B = - 2.003): Este valor establece que cuando el género es femenino, la mortalidad cae en 2.003, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 1.101): Este valor aclara que, dado que hay presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 1.101, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 21.249): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 21.249, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 3.231): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 3.231 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.865): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.865, manteniendo constantes las otras condiciones.

**ID\_G00\_G99** (B = 1.359): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 1.359 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_I00\_I99** (B = 25.122): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 25.122, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 9.920): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 9.920 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 6.253): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 6.253, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.163): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.163 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 1.787): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 1.787, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 1.182): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.182 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 1.387): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 1.387, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.858): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.858 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $B = 2.639$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificadas en otra parte, la mortalidad permanece en 2.639, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $B = 8.401$ ): Este valor establece que, como existe la presencia de lesiones, envenenamientos y otras consecuencias de causas externas, generará un aumento de 8.401 en la mortalidad, manteniendo constantes las otras condiciones.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.073$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.04), la mortalidad aumenta en un 7.3% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.699 ( $0.073 * 9.580$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.020$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.43), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.191 ( $0.020 * 9.580$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.082$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 8.2% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.785 ( $0.082 * 9.580$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.105$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 10.5% de

las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 1.005 ( $0.105 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.028$ ): Este valor aclara que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.245), la mortalidad aumenta en un 2.8% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.268 ( $0.028 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.612$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.276), la mortalidad aumenta en un 61.2% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 5.862 ( $0.612 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.092$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.273), la mortalidad aumenta en un 9.2% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.881 ( $0.092 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.022$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.246), la mortalidad aumenta en un 2.2% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.210 ( $0.022 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.037$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.257), la mortalidad aumenta en un 3.7% de las desviaciones estándar. La desviación

estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.354 ( $0.037 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.724$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.276), la mortalidad aumenta en un 72.4% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 6.935 ( $0.724 * 9.580$ ), manteniendo constantes las demás condiciones.

**ID\_J00\_J99** ( $\beta = 0.286$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.276), la mortalidad aumenta en un 28.6% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 2.739 ( $0.286 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.178$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.273), la mortalidad aumenta en un 17.8% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 1.705 ( $0.178 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.003$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.168), la mortalidad aumenta en un 0.3% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.028 ( $0.003 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.048$ ): Este valor establece que cuando hay una presencia de enfermedades del sistema genitourinario con una desviación estándar (0.258), la mortalidad aumenta en un 4.8% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.459 ( $0.048 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.009$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.074), la mortalidad aumenta en un 0.9% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.086 ( $0.009 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.027$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.183), la mortalidad aumenta en un 2.7% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.258 ( $0.027 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.018$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.199), la mortalidad aumenta en un 1.8% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.172 ( $0.018 * 9.580$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.068$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.246), la mortalidad aumenta en un 6.8% de desviaciones estándar. La desviación estándar del número de muertes es de 9.580, por lo que esto constituye un cambio de 0.651 ( $0.068 * 9.580$ ), manteniendo constantes las demás condiciones.

**ID\_S00\_T98** ( $\beta = 0.24$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.273), la mortalidad aumenta en un 24% de las desviaciones estándar. La desviación estándar del número de muertes es de



9.580, por lo que esto constituye un cambio de 2.299 ( $0.24 * 9.580$ ), manteniendo las otras condiciones constantes.

Tabla 112: *Evaluación - Resumen del modelo de RLM de Araucanía*

<b>Model</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Araucanía model	0,702	0,701	6,30366	1,871

Se puede ver que el modelo generado para la región de la Araucanía presenta un R cuadrado ajustado de 0.701, es decir que la varianza de la mortalidad se explica al 70.1% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 113: *Evaluación - ANOVA del modelo de RLM de Araucanía*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	1179013.501	58950.675	1483.553	.000
Residue	501152.110	39.736	-	-
Total	1680165.611	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 114: *Evaluación - Estadísticos descriptivos del modelo de Araucanía*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	9.6092	11.53294
TYear	2007.2002	6.05917
TMonth	6.5320	3.43389
DLocationID	.5286	.49920
SexID	.5039	.50000
ID_A00_B99	.0712	.25724
ID_C00_D48	.0798	.27098
ID_E00_E90	.0785	.26901
ID_F00_F99	.0660	.24832
ID_G00_G99	.0718	.25816
ID_I00_I99	.0798	.27098
ID_J00_J99	.0798	.27098
ID_K00_K93	.0780	.26814
ID_L00_L99	.0282	.16549
ID_M00_M99	.0333	.17949
ID_N00_N99	.0721	.25868
ID_O00_O99	.0051	.07100
ID_P00_P96	.0378	.19062
ID_Q00_Q99	.0457	.20879
ID_R00_R99	.0634	.24370
ID_S00_T98	.0786	.26913

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 9.6092.

Tabla 115: *Evaluación - Coeficientes del modelo de RLM de Araucanía*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-205.497	18.609	-	-11.043	.000	-
TYear	.104	.009	.055	11.240	.000	1.002
TMonth	.083	.016	.025	5.072	.000	1.001
DLocationID	-3.334	.115	-.144	-29.065	.000	1.042
SexID	-1.844	.113	-.080	-16.351	.000	1.010
ID_A00_B99	1.557	.382	.035	4.077	.000	3.068
ID_C00_D48	27.096	.376	.637	72.109	.000	3.296

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_E00_E90	3.343	.377	.078	8.876	.000	3.262
ID_F00_F99	1.182	.387	.025	3.056	.002	2.932
ID_G00_G99	1.885	.382	.042	4.939	.000	3.085
ID_I00_I99	28.759	.376	.676	76.537	.000	3.296
ID_J00_J99	10.860	.376	.255	28.900	.000	3.296
ID_K00_K93	6.047	.377	.141	16.042	.000	3.247
ID_L00_L99	.126	.462	.002	.272	.785	1.857
ID_M00_M99	.394	.443	.006	.889	.374	2.012
ID_N00_N99	2.072	.381	.046	5.435	.000	3.091
ID_O00_O99	1.444	.852	.009	1.694	.090	1.164
ID_P00_P96	1.876	.431	.031	4.349	.000	2.149
ID_Q00_Q99	1.279	.413	.023	3.094	.002	2.368
ID_R00_R99	9.806	.390	.207	25.159	.000	2.868
ID_S00_T98	10.917	.377	.255	28.990	.000	3.265

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 205.497 + 0.10(\text{TYear}) + 0.08(\text{TMonth}) - 3.33 (\text{DLocationID}) - 1.84(\text{SexID}) \\ & + 1.55(\text{ID\_A00\_B99}) + 27.09(\text{ID\_C00\_D48}) + 3.34(\text{ID\_E00\_E90}) + 1.18(\text{ID\_F00\_F99}) + \\ & 1.88(\text{ID\_G00\_G99}) + 28.75(\text{ID\_I00\_I99}) + 10.86(\text{ID\_J00\_J99}) + 6.04(\text{ID\_K00\_K93}) - \\ & 0.(\text{ID\_L00\_L99}) + 0.39(\text{ID\_M00\_M99}) + 2.07(\text{ID\_N00\_N99}) + 1.44(\text{ID\_O00\_O99}) + \\ & 1.87(\text{ID\_P00\_P96}) + 1.27(\text{ID\_Q00\_Q99}) + 9.80(\text{ID\_R00\_R99}) + 10.917(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.104): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.104 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.083): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de mortalidad de 0.083, manteniendo las otras condiciones constantes.

**DLocationID** (B = -3.334): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 3.334, manteniendo las otras condiciones constantes.

**SexID** (B = - 1.844): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.844, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 1.557): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 1.557, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 27.096): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 27.096, manteniendo constantes las otras condiciones.

**ID\_E00\_E90** (B = 3.343): Este valor indica que, dado que existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 3.343 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 1.182): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 1.182, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 1.885): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 1.885 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_I00\_I99** (B = 28.759): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 28.759, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 10.860): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 10.860 en la mortalidad, manteniendo constantes las demás afecciones.

**ID\_K00\_K93** (B = 6.047): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 6.047, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.394): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.394 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 2.072): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 2.072, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 1.444): Este valor aclara que, dado que existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.444 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 1.876): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 1.876, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 1.279): Este valor indica que debido a la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.279 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_R00\_R99** ( $B = 9.806$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 9.806, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $B = 10.917$ ): Este valor establece que, como existe la presencia de lesiones, envenenamientos y algunas otras consecuencias de causas externas, generará un aumento de 10.917 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.055$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.06), la mortalidad aumenta en 5.5% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.634 ( $0.055 * 11.532$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.025$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.43), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.283 ( $0.025 * 11.532$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.144$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 14.4% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 1.66 ( $0.144 * 11.532$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.080$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 8% de las desviaciones estándar. La desviación estándar del número de muertes es de

11.532, por lo que esto constituye un cambio de 0.922 ( $0.080 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.035$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.257), la mortalidad aumenta en un 3.5% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.403 ( $0.035 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.637$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.270), la mortalidad aumenta en un 63.7% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 7.345 ( $0.637 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.078$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.269), la mortalidad aumenta en un 7.8% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.899 ( $0.078 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.025$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.248), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 11.532, por lo que esto constituye un cambio de 0.288 ( $0.025 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.042$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.258), la mortalidad aumenta en un 4.2% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un

cambio de 0.484 (0.042 \* 11.532), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.676$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.270), la mortalidad aumenta en un 67.6% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 7.795 (0.676 \* 11.532), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.255$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.270), la mortalidad aumenta en un 25.5% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 2.940 (0.255 \* 11.532), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.141$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.268), la mortalidad aumenta en un 14.1% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 1.626 (0.141 \* 11.532), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.006$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.179), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 11.532, por lo que esto constituye un cambio de 0.07 (0.006 \* 11.532), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.046$ ): Este valor establece que cuando hay una presencia de enfermedades del sistema genitourinario con una desviación estándar (0.258), la mortalidad aumenta en un 4.6% de las desviaciones estándar. La desviación



estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.53 ( $0.046 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.009$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.071), la mortalidad aumenta en un 0.9% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.10 ( $0.009 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.031$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.190), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.357 ( $0.031 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.023$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.208), la mortalidad aumenta en un 2.3% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 0.265 ( $0.023 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.207$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.243), la mortalidad aumenta en un 20.7% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 2.387 ( $0.207 * 11.532$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.255$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.269), la mortalidad aumenta en un 25.5% de las desviaciones estándar. La desviación estándar del número de muertes es de 11.532, por lo que esto constituye un cambio de 2.940 ( $0.255 * 11.532$ ), manteniendo las otras condiciones constantes.

Tabla 116: *Evaluación - Resumen del modelo de RLM de Coquimbo*

<b>Modelo</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Coquimbo	0.699	0.698	4.26049	2.089

Se puede ver que el modelo generado para la región de Coquimbo presenta una R cuadrado ajustado de 0.698, es decir que la varianza de la mortalidad se explica al 69.8% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 117: *Evaluación - ANOVA del modelo de RLM de Coquimbo*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	481172.696	24058.635	1325.413	.000
Residue	207348.022	18.152	-	-
Total	688520.718	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 118: *Evaluación - Estadísticos descriptivos del modelo de Coquimbo*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	6.4194	7.75691
TYear	2007.3121	6.02511
TMonth	6.5375	3.42759
DLocationID	.5147	.49981
SexID	.4996	.50002
ID_A00_B99	.0744	.26237
ID_C00_D48	.0881	.28343
ID_E00_E90	.0825	.27512
ID_F00_F99	.0509	.21989
ID_G00_G99	.0719	.25836
ID_I00_I99	.0881	.28343
ID_J00_J99	.0879	.28317
ID_K00_K93	.0837	.27697
ID_L00_L99	.0188	.13578
ID_M00_M99	.0289	.16760
ID_N00_N99	.0737	.26123
ID_O00_O99	.0050	.07040
ID_P00_P96	.0339	.18099
ID_Q00_Q99	.0420	.20067
ID_R00_R99	.0632	.24329
ID_S00_T98	.0831	.27605

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 6.4194

Tabla 119: *Evaluación - Coeficientes del modelo de RLM de Coquimbo*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-264.701	13.309		-19.889	.000	
TYear	.133	.007	.103	20.113	.000	1.004
TMonth	.046	.012	.020	3.940	.000	1.001
DLocationID	-2.511	.082	-.162	-30.497	.000	1.068
SexID	-1.057	.080	-.068	-13.202	.000	1.010
ID_A00_B99	1.353	.296	.046	4.570	.000	3.805
ID_C00_D48	18.113	.291	.662	62.344	.000	4.275

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_E00_E90	2.268	.293	.080	7.752	.000	4.085
ID_F00_F99	.158	.313	.004	.507	.612	2.980
ID_G00_G99	1.010	.297	.034	3.397	.001	3.721
ID_I00_I99	18.478	.291	.675	63.601	.000	4.275
ID_J00_J99	6.088	.291	.222	20.951	.000	4.269
ID_K00_K93	3.599	.292	.129	12.322	.000	4.125
ID_L00_L99	-.297	.389	-.005	-.764	.445	1.757
ID_M00_M99	.130	.348	.003	.374	.709	2.147
ID_N00_N99	1.122	.296	.038	3.785	.000	3.779
ID_O00_O99	1.356	.622	.012	2.179	.029	1.210
ID_P00_P96	1.876	.338	.044	5.556	.000	2.354
ID_Q00_Q99	1.316	.323	.034	4.073	.000	2.650
ID_R00_R99	2.796	.303	.088	9.220	.000	3.432
ID_S00_T98	4.399	.292	.157	15.044	.000	4.107

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 264.70 + 0.13(\text{TYear}) + 0.46(\text{TMonth}) - 2.51 (\text{DLocationID}) - 1.05(\text{SexID}) \\ & + 1.35(\text{ID\_A00\_B99}) + 18.11(\text{ID\_C00\_D48}) + 2.26(\text{ID\_E00\_E90}) + 0.15(\text{ID\_F00\_F99}) + \\ & 1.01(\text{ID\_G00\_G99}) + 18.47(\text{ID\_I00\_I99}) + 6.08(\text{ID\_J00\_J99}) + 3.59(\text{ID\_K00\_K93}) - \\ & 0.29(\text{ID\_L00\_L99}) + 0.13(\text{ID\_M00\_M99}) + 1.12(\text{ID\_N00\_N99}) + 1.35(\text{ID\_O00\_O99}) + \\ & 1.87(\text{ID\_P00\_P96}) + 1.31(\text{ID\_Q00\_Q99}) + 2.79(\text{ID\_R00\_R99}) + 4.39(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.133): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.133 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.046): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.046 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -2.511): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 2.511, manteniendo las otras condiciones constantes.

**SexID** (B = -1.057): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.057, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 1.353): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.353, sufren las otras condiciones constantes.

**ID\_C00\_D48** (B= 18.113): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 18.113, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 2.268): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 2.268 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.158): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.158, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 1.010): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 1.010 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 18.478): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 18.478, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 6.088): Este valor indica que, dado que hay presencia de enfermedades del sistema respiratorio, generará un aumento de 6.088 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 3.599): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 3.599, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.130): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.130 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 1.122): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generar un aumento de 1.122 en la mortalidad, generar las otras condiciones constantes.

**ID\_O00\_O99** (B = 1.356): Este valor aclara que como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.356 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 1.876): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 1.876, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 1.316): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.316 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $B = 2.796$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 2.796, manteniendo constantes las otras condiciones.

**ID\_S00\_T98** ( $B = 4.399$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 4.399 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.103$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.02), la mortalidad aumenta en un 10.3% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.798 ( $0.103 * 7.756$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.020$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.42), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.155 ( $0.020 * 7.756$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.162$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 16.2% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 1.256 ( $0.162 * 7.756$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.068$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae un 6.8% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756,

por lo que esto constituye un cambio de 0.527 ( $0.068 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.046$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.262), la mortalidad aumenta en un 4.6% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.356 ( $0.046 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.662$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.283), la mortalidad aumenta en un 66.2% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 5.13 ( $0.662 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.080$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.275), la mortalidad aumenta en un 8% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 3.620 ( $0.080 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.004$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.219), la mortalidad aumenta en un 0.4% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.03 ( $0.004 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.034$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.258), la mortalidad aumenta en un 3.4% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.263 ( $0.034 * 7.756$ ), manteniendo las otras condiciones constantes.



**ID\_I00\_I99** ( $\beta = 0.675$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.283), la mortalidad aumenta en un 67.5% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 5.235 ( $0.675 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.222$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0,283), la mortalidad aumenta en un 22.2% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 1.721 ( $0.222 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.129$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.276), la mortalidad aumenta en un 12.9% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 1 ( $0.129 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.003$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.167), la mortalidad aumenta en un 0.3% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.02 ( $0.003 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.038$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.261), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.294 ( $0.038 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.012$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar

(0.070), la mortalidad aumenta en un 1.2% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.093 ( $0.012 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.044$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.180), la mortalidad aumenta en un 4.4% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.341 ( $0.044 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.034$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.2), la mortalidad aumenta en un 3.4% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.263 ( $0.034 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.088$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.243), la mortalidad aumenta en 8.8% desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 0.682 ( $0.088 * 7.756$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.157$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.276), la mortalidad aumenta en un 15.7% de las desviaciones estándar. La desviación estándar del número de muertes es 7.756, por lo que esto constituye un cambio de 1.217 ( $0.157 * 7.756$ ), manteniendo las otras condiciones constantes.

Tabla 120: *Evaluación - Resumen del modelo de RLM de Ñuble*

<b>Modelo</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Ñuble	0.690	0.689	3.73417	2.015

Se puede ver que el modelo generado para la región de Ñuble presenta un R cuadrado ajustado de 0.689, es decir que la varianza de la mortalidad se explica al 68.9% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 121: *Evaluación - ANOVA del modelo de RLM de Ñuble*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	333399.982	16669.999	1195.497	.000
Residue	150107.083	13.944	-	-
Total	483507.066	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 122: *Evaluación - Estadísticos descriptivos del modelo de Ñuble*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	5.9161	6.69563
TYear	2007.1173	6.04541
TMonth	6.5141	3.42667
DLocationID	.5261	.49934
SexID	.4977	.50002
ID_A00_B99	.0654	.24718
ID_C00_D48	.0935	.29108

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
ID_E00_E90	.0855	.27961
ID_F00_F99	.0607	.23884
ID_G00_G99	.0704	.25578
ID_I00_I99	.0935	.29108
ID_J00_J99	.0922	.28926
ID_K00_K93	.0888	.28450
ID_L00_L99	.0223	.14751
ID_M00_M99	.0229	.14959
ID_N00_N99	.0706	.25609
ID_O00_O99	.0025	.04997
ID_P00_P96	.0280	.16498
ID_Q00_Q99	.0350	.18367
ID_R00_R99	.0547	.22741
ID_S00_T98	.0877	.28288

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 5.916

Tabla 123: *Evaluación - Coeficientes del modelo de RLM de Ñuble*

<b>Región</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-107.672	11.967		-8.997	.000	
TYear	.055	.006	.049	9.200	.000	1.004
TMonth	.043	.011	.022	4.063	.000	1.001
DLocationID	-1.217	.075	-.091	-16.331	.000	1.072
SexID	-1.344	.072	-.100	-18.607	.000	1.009
ID_A00_B99	.784	.262	.029	2.998	.003	3.233
ID_C00_D48	13.384	.250	.582	53.510	.000	4.100
ID_E00_E90	1.947	.253	.081	7.708	.000	3.858

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_F00_F99	.697	.265	.025	2.631	.009	3.099
ID_G00_G99	.614	.259	.023	2.369	.018	3.395
ID_I00_I99	17.134	.250	.745	68.503	.000	4.100
ID_J00_J99	4.700	.251	.203	18.762	.000	4.061
ID_K00_K93	4.313	.252	.183	17.149	.000	3.960
ID_L00_L99	-.115	.327	-.003	-.351	.726	1.797
ID_M00_M99	.178	.325	.004	.547	.584	1.823
ID_N00_N99	1.005	.259	.038	3.883	.000	3.396
ID_O00_O99	.892	.753	.007	1.185	.236	1.095
ID_P00_P96	.522	.309	.013	1.688	.091	2.009
ID_Q00_Q99	.559	.293	.015	1.909	.056	2.241
ID_R00_R99	.979	.271	.033	3.619	.000	2.928
ID_S00_T98	5.227	.252	.221	20.742	.000	3.930

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 107.672 + 0.05(\text{TYear}) + 0.43(\text{TMonth}) - 1.12 (\text{DLocationID}) - 1.34(\text{SexID}) \\ & + 0.78(\text{ID\_A00\_B99}) + 13.38(\text{ID\_C00\_D48}) + 1.94(\text{ID\_E00\_E90}) + 0.69(\text{ID\_F00\_F99}) + \\ & 0.61(\text{ID\_G00\_G99}) + 17.13(\text{ID\_I00\_I99}) + 4.70(\text{ID\_J00\_J99}) + 4.31(\text{ID\_K00\_K93}) - \\ & 0.11(\text{ID\_L00\_L99}) + 0.17(\text{ID\_M00\_M99}) + 1(\text{ID\_N00\_N99}) + 0.89(\text{ID\_O00\_O99}) + \\ & 0.52(\text{ID\_P00\_P96}) + 0.55(\text{ID\_Q00\_Q99}) + 0.97(\text{ID\_R00\_R99}) + 5.22(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.055): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.055 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.043): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.043 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -1.217): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 1.217, manteniendo las otras condiciones constantes.

**SexID** (B = -1.344): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.344, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.784): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.784, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 13.384): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 13.384, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 1.947): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 1.947 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.697): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.697, manteniendo constantes las otras condiciones.

**ID\_G00\_G99** (B = 0.614): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.614 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 17.134): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 17.134, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 4.7): Este valor indica que, como existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 4.7 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_K00\_K93** (B = 4.313): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 4.313, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.178): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.178 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 1.005): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 1.005, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.892): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.892 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 0.522): Este valor indica que cuando hay presencia en ciertas condiciones que se originan en el período perinatal, la mortalidad permanece en 0.522, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.559): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.559 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 0.979): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.979, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 5.227$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y otras consecuencias de causas externas, generará un aumento de 5.227 en la mortalidad, manteniendo constantes las otras condiciones.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.049$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.04), la mortalidad aumenta en un 4.9% de desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.328 ( $0.049 * 6.695$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.022$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.42), la mortalidad aumenta en un 2.2% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.133 ( $0.02 * 6.695$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.091$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae un 9.1% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.609 ( $0.091 * 6.695$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.100$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 10% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.669 ( $0.10 * 6.695$ ), manteniendo las otras condiciones constantes.



**ID\_A00\_B99** ( $\beta = 0.029$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.247), la mortalidad aumenta en un 2.9% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.194 ( $0.029 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.582$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una con una desviación estándar (0.291), la mortalidad aumenta en un 58.9% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 3.896 ( $0.582 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.081$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.279), la mortalidad aumenta en un 8.1% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.542 ( $0.081 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.025$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.238), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.167 ( $0.025 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.023$ ): Este valor establece que, dado que hay enfermedades del sistema nervioso central con una desviación estándar (0.255), la mortalidad aumenta en un 2.3% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 1.153 ( $0.023 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.745$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.291), la

mortalidad aumenta en un 74.5% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 4.987 ( $0.745 * 6.695$ ), manteniendo las otras condiciones constantes

**ID\_J00\_J99** ( $\beta = 0.203$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.289), la mortalidad aumenta en un 20.3% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 1.359 ( $0.203 * 6.695$ ), manteniendo constantes las otras condiciones.

**ID\_K00\_K93** ( $\beta = 0.183$ ): Este valor indica que cuando hay una presencia de enfermedades del sistema digestivo con una desviación estándar (0.284), la mortalidad aumenta en un 18.3% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 1.225 ( $0.183 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.004$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.149), la mortalidad aumenta en un 0.4% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.026 ( $0.004 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.038$ ): Este valor establece que cuando hay una presencia de enfermedades del sistema genitourinario con una desviación estándar (0.256), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.254 ( $0.038 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.007$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.049), la mortalidad aumenta en un 0.7% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye

un cambio de 0.04 ( $0.007 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.013$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.164), la mortalidad aumenta en un 1.3% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.087 ( $0.013 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.015$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.183), la mortalidad aumenta en un 1.5% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.308 ( $0.015 * 6.695$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.033$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.227), la mortalidad aumenta en un 3.3% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 0.1 ( $0.033 * 6.695$ ), manteniendo las otras condiciones constantes

**ID\_S00\_T98** ( $\beta = 0.221$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.282), la mortalidad aumenta en un 22.1% de las desviaciones estándar. La desviación estándar del número de muertes es 6.695, por lo que esto constituye un cambio de 1.479 ( $0.221 * 6.695$ ), manteniendo las otras condiciones constantes.

Tabla 124: *Evaluación - Resumen del modelo de RLM de Antofagasta*

<b>Model</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Antofagasta	0.679	0.678	3.23066	2.022

Se puede observar que el modelo generado para la región de Antofagasta presenta un R cuadrado ajustado de 0.678, es decir que la varianza de la mortalidad se explica al 67.8% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 125: *Evaluación - ANOVA del modelo de RLM de Antofagasta*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	235801.514	11790.076	1129.627	.000
Residue	111583.427	10.437	-	-
Total	347384.941	=	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 126: *Evaluación - Estadísticos descriptivos del modelo de Antofagasta*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	5,1459	5,69496
TYear	2007,2508	6,06145
TMonth	6,5390	3,42089
DLocationID	,5437	,49811
SexID	,5025	,50002
ID_A00_B99	,0595	,23651
ID_C00_D48	,0941	,29198
ID_E00_E90	,0837	,27701
ID_F00_F99	,0611	,23944
ID_G00_G99	,0635	,24384
ID_I00_I99	,0941	,29198
ID_J00_J99	,0925	,28976

Variable	Mean	Standard Deviation
ID_K00_K93	,0851	,27910
ID_L00_L99	,0164	,12713
ID_M00_M99	,0244	,15419
ID_N00_N99	,0724	,25923
ID_O00_O99	,0047	,06816
ID_P00_P96	,0380	,19119
ID_Q00_Q99	,0400	,19608
ID_R00_R99	,0612	,23978
ID_S00_T98	,0879	,28322

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 5.1459.

Tabla 127: *Evaluación - Coeficientes del modelo de RLM de Antofagasta*

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-145.089	10.358		-14.007	.000	
TYear	.073	.005	.078	14.200	.000	1.003
TMonth	.029	.009	.018	3.222	.001	1.000
DLocationID	-.612	.065	-.054	-9.411	.000	1.077
SexID	-1.175	.063	-.103	-18.722	.000	1.011
ID_A00_B99	.885	.249	.037	3.548	.000	3.570
ID_C00_D48	14.236	.237	.730	60.000	.000	4.925
ID_E00_E90	1.744	.240	.085	7.270	.000	4.530
ID_F00_F99	.962	.249	.040	3.860	.000	3.658
ID_G00_G99	.754	.248	.032	3.047	.002	3.739
ID_I00_I99	12.057	.237	.618	50.818	.000	4.925
ID_J00_J99	3.809	.238	.194	16.030	.000	4.865
ID_K00_K93	3.170	.239	.155	13.243	.000	4.582
ID_L00_L99	.145	.324	.003	.447	.655	1.745
ID_M00_M99	.226	.293	.006	.772	.440	2.096

Variable	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_N00_N99	1.175	.243	.053	4.825	.000	4.088
ID_O00_O99	.648	.506	.008	1.281	.200	1.219
ID_P00_P96	.997	.268	.033	3.721	.000	2.690
ID_Q00_Q99	.720	.265	.025	2.717	.007	2.772
ID_R00_R99	.720	.249	.030	2.886	.004	3.672
ID_S00_T98	3.886	.239	.193	16.271	.000	4.695

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ\text{Deaths} = & - 145.08 + 0.07(\text{TYear}) + 0.029(\text{TMonth}) - 0.61 (\text{DLocationID}) - 1.17(\text{SexID}) \\ & + 0.88(\text{ID\_A00\_B99}) + 14.23(\text{ID\_C00\_D48}) + 1.74(\text{ID\_E00\_E90}) + 0.96(\text{ID\_F00\_F99}) + \\ & 0.75(\text{ID\_G00\_G99}) + 12.05(\text{ID\_I00\_I99}) + 3.80(\text{ID\_J00\_J99}) + 3.17(\text{ID\_K00\_K93}) - \\ & 0.145(\text{ID\_L00\_L99}) + 0.22(\text{ID\_M00\_M99}) + 1.17(\text{ID\_N00\_N99}) + 0.64(\text{ID\_O00\_O99}) + \\ & 0.99(\text{ID\_P00\_P96}) + 0.72(\text{ID\_Q00\_Q99}) + 0.72(\text{ID\_R00\_R99}) + 3.88(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.073): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.073 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.029): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de la mortalidad de 0.029, manteniendo las otras condiciones constantes.

**DLocationID** (B = -0.612): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.612, manteniendo las otras condiciones constantes.

**SexID** (B = -1.175): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.175, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.885): Este valor aclara que, dado que hay presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.885, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 14.236): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 14.236, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 1.744): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 1.744 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.962): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.962 manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.754): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.754 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 12.057): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 12.057, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 3.809): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 3.809 en la mortalidad, manteniendo constantes las otras condiciones

**ID\_K00\_K93** (B = 3.170): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 3.170, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.226): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.226 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 1.175): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 1.175, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.648): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.648 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 0.997): Este valor indica que cuando hay presencia en ciertas condiciones que se originan en el período perinatal, la mortalidad permanece en 0.997, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.720): Este valor indica que debido a la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.720 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 0.720): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.720, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 3.886): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 3.886 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:



**Tyear** ( $\beta = 0.078$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.06), la mortalidad aumenta en un 7.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.444 ( $0.078 * 5.694$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.018$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.42), la mortalidad aumenta en un 1.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.102 ( $0.018 * 5.694$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.054$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae un 5.4% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.307 ( $0.054 * 5.694$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.103$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 10.3% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.586 ( $0.103 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.037$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.236), la mortalidad aumenta en un 3.7% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.210 ( $0.037 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.730$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.291), la mortalidad aumenta en un 73% de las desviaciones estándar. La desviación estándar del

número de muertes es 5.694, por lo que esto constituye un cambio de 4.156 ( $0.730 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.085$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.277), la mortalidad aumenta en un 8.5% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.483 ( $0.085 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.040$ ): Este valor determina que cuando hay presencia de enfermedades mentales y de trastornos del comportamiento con una desviación estándar (0.239), la mortalidad aumenta en un 4% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.227 ( $0.040 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.032$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.243), la mortalidad aumenta en un 3.2% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.182 ( $0.032 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.618$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.291), la mortalidad aumenta en un 61.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 3.518 ( $0.618 * 5.694$ ), manteniendo constantes las otras condiciones.

**ID\_J00\_J99** ( $\beta = 0.194$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.289), la mortalidad aumenta en un 19.4% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 1.104 ( $0.194 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.155$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.279), la mortalidad aumenta en un 15.5% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.882 ( $0.155 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.006$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.154), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.034 ( $0.006 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.053$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.259), la mortalidad aumenta en un 5.3% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.301 ( $0.053 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.008$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.068), la mortalidad aumenta en un 0.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.04 ( $0.008 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.033$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.068), la mortalidad aumenta en un 0.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.04 ( $0.008 * 5.694$ ), las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.025$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.142), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.142 ( $0.025 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.030$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.239), la mortalidad aumenta en un 3% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 0.170 ( $0.030 * 5.694$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.193$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.283), la mortalidad aumenta en un 19.3% de las desviaciones estándar. La desviación estándar del número de muertes es 5.694, por lo que esto constituye un cambio de 1.098 ( $0.193 * 5.694$ ), manteniendo las otras condiciones constantes.

Tabla 128: *Evaluación - Resumen del modelo de RLM de Los Lagos*

Modelo	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Los Lagos	0.678	0.677	5.32778	2.027

Se puede ver que el modelo generado para la región de Los Lagos presenta un R cuadrado ajustado de 0.677, es decir que la varianza de la mortalidad se explica al 67.7% por las variables independientes. Por otro lado, en la prueba de la estadística de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 129: *Evaluación - ANOVA del modelo de RLM de Los Lagos*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	721623.626	36081.181	1271.127	.000
Residue	343489.244	28.385	-	-
Total	1065112.869	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 130: *Evaluación - Estadísticos descriptivos del modelo de Los Lagos*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	7,9757	9,37408
TYear	2007,2139	6,04599
TMonth	6,5306	3,43256
DLocationID	,5177	,49971
SexID	,5001	,50002
ID_A00_B99	,0701	,25536
ID_C00_D48	,0832	,27613
ID_E00_E90	,0808	,27261
ID_F00_F99	,0604	,23821
ID_G00_G99	,0725	,25935
ID_I00_I99	,0832	,27613
ID_J00_J99	,0832	,27613
ID_K00_K93	,0819	,27425
ID_L00_L99	,0274	,16322
ID_M00_M99	,0323	,17668
ID_N00_N99	,0733	,26070
ID_O00_O99	,0032	,05663
ID_P00_P96	,0370	,18887
ID_Q00_Q99	,0426	,20189
ID_R00_R99	,0610	,23942
ID_S00_T98	.0799	.27121

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es de 7.9757.

Tabla 131: *Evaluación - Coeficientes del modelo de RLM de Los Lagos*

<b>Región</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-211.637	16.099		-13.146	.000	
TYear	.107	.008	.069	13.377	.000	1.003
TMonth	.094	.014	.035	6.691	.000	1.000
DLocationID	-3.477	.100	-.185	-34.918	.000	1.057
SexID	-2.002	.097	-.107	-20.577	.000	1.011
ID_A00_B99	1.357	.342	.037	3.964	.000	3.265
ID_C00_D48	21.246	.335	.626	63.463	.000	3.649
ID_E00_E90	2.870	.336	.083	8.544	.000	3.581
ID_F00_F99	.539	.351	.014	1.537	.124	2.979
ID_G00_G99	1.364	.341	.038	4.002	.000	3.337
ID_I00_I99	22.281	.335	.656	66.557	.000	3.649
ID_J00_J99	8.631	.335	.254	25.783	.000	3.649
ID_K00_K93	5.773	.335	.169	17.214	.000	3.613
ID_L00_L99	-.352	.412	-.006	-.856	.392	1.927
ID_M00_M99	.519	.396	.010	1.312	.189	2.088
ID_N00_N99	1.396	.340	.039	4.103	.000	3.360
ID_O00_O99	1.761	.903	.011	1.951	.051	1.116
ID_P00_P96	2.030	.385	.041	5.273	.000	2.259
ID_Q00_Q99	1.553	.373	.033	4.159	.000	2.426
ID_R00_R99	5.086	.350	.130	14.513	.000	3.006
ID_S00_T98	9.381	.337	.271	27.877	.000	3.557

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ\text{Deaths} = & - 211.63 + 0.10(\text{TYear}) + 0.09(\text{TMonth}) - 3.47 (\text{DLocationID}) - 2(\text{SexID}) + \\ & 1.35(\text{ID\_A00\_B99}) + 21.24(\text{ID\_C00\_D48}) + 2.87(\text{ID\_E00\_E90}) + 0.53(\text{ID\_F00\_F99}) + \\ & 1.36(\text{ID\_G00\_G99}) + 22.28(\text{ID\_I00\_I99}) + 8.63(\text{ID\_J00\_J99}) + 5.77(\text{ID\_K00\_K93}) - \\ & 0.35(\text{ID\_L00\_L99}) + 0.51(\text{ID\_M00\_M99}) + 1.39(\text{ID\_N00\_N99}) + 1.76(\text{ID\_O00\_O99}) + \\ & 2.03(\text{ID\_P00\_P96}) + 1.55(\text{ID\_Q00\_Q99}) + 5.08(\text{ID\_R00\_R99}) + 9.38(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.107): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.107 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.094): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.094 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -3.477): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 3.477, manteniendo las otras condiciones constantes.

**SexID** (B = -2.002): Este valor establece que cuando el género es femenino, la mortalidad cae en 2.002, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 1.357): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.357, tienen las otras condiciones constantes.

**ID\_C00\_D48** (B= 21.246): Este valor establece que cuando el género es femenino, la mortalidad cae en 21.246, tienen las otras condiciones constantes.

**ID\_E00\_E90** (B = 2.870): Este valor indica que, dado que existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 2.870 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.539): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.539, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 1.364): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 1.364 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 22.281): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 22.281, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 8.631): Este valor indica que, dado que hay presencia de enfermedades del sistema respiratorio, generará un aumento de 8.631 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 5.773): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 5.773, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.519): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.519 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 1.396): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 1.396, manteniendo constantes las otras condiciones.

**ID\_O00\_O99** (B = 1.761): Este valor aclara que, dado que existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 1.761 en la mortalidad, manteniendo constantes las otras condiciones.



**ID\_P00\_P96** ( $B = 2.030$ ): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 2.030, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $B = 1.553$ ): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 1.553 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $B = 5.086$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 5.086, manteniendo constantes las otras condiciones.

**ID\_S00\_T98** ( $B = 9.381$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 9.381 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.069$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.04), la mortalidad aumenta en un 6.9% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.646 ( $0.069 * 9.374$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.035$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.43), la mortalidad aumenta en un 3.5% de desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.328 ( $0.035 * 9.374$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.185$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 18.5% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 1.734 ( $0.185 * 9.374$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.107$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 10.7% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 1.003 ( $0.107 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.037$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.255), la mortalidad aumenta en un 3.7% de las desviaciones estándar. La desviación estándar del número de muertes es de 9.374, por lo que esto constituye un cambio de 0.346 ( $0.037 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.626$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.276), la mortalidad aumenta en un 62.6% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 5.868 ( $0.626 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.083$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.272), la mortalidad aumenta en un 8.3% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.778 ( $0.083 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.014$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.238), la mortalidad aumenta en un 1.4% de las desviaciones

estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.131 ( $0.014 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.038$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.259), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.356 ( $0.038 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.656$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.276), la mortalidad aumenta en un 65.6% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 6.149 ( $0.656 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.254$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.276), la mortalidad aumenta en un 25.4% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 2.38 ( $0.254 * 9.374$ ), manteniendo constantes las otras condiciones.

**ID\_K00\_K93** ( $\beta = 0.169$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.274), la mortalidad aumenta en un 16.9% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 1.584 ( $0.169 * 9.374$ ), manteniendo constantes las otras condiciones.

**ID\_M00\_M99** ( $\beta = 0.010$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.176), la mortalidad aumenta en un 1% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.093 ( $0.010 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.039$ ): Este valor establece que cuando hay una presencia de enfermedades del sistema genitourinario con una desviación estándar (0.260), la mortalidad aumenta en un 3.9% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.365 ( $0.039 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.011$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.056), la mortalidad aumenta en un 1.1% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.103 ( $0.011 * 9.374$ ), manteniendo constantes las otras condiciones.

**ID\_P00\_P96** ( $\beta = 0.041$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.188), la mortalidad aumenta en un 4.1% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.384 ( $0.041 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.033$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.201), la mortalidad aumenta en un 3.3% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 0.309 ( $0.033 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.130$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.239), la mortalidad aumenta en un 13% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 1.218 ( $0.130 * 9.374$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.271$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.271), la mortalidad aumenta en un 27.1% de las desviaciones estándar. La desviación estándar del número de muertes es 9.374, por lo que esto constituye un cambio de 2.54 ( $0.271 * 9.374$ ), manteniendo las otras condiciones constantes.

Tabla 132: *Evaluación - Resumen del modelo de RLM de Los Ríos*

<b>Modelo</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Los Ríos modelo	0.670	0.669	2.97409	1.967

Se puede ver que el modelo generado para la región de Los Ríos presenta un R cuadrado ajustado de 0.669, es decir que la varianza de la mortalidad se explica al 66.9% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 133: *Evaluación - ANOVA del modelo de RLM de Los Ríos*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	183321.015	9166.051	1036.273	.000
Residue	90415.713	8.845	-	-
Total	273736.728	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 134: *Evaluación - Estadísticos descriptivos del modelo de Los Ríos*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	4.8930	5.16980
TYear	2007.3048	6.03392
TMonth	6.5409	3.43324
DLocationID	.5220	.49954
SexID	.4909	.49994
ID_A00_B99	.0624	.24186
ID_C00_D48	.0984	.29788
ID_E00_E90	.0840	.27734
ID_F00_F99	.0577	.23318
ID_G00_G99	.0695	.25433
ID_I00_I99	.0984	.29788
ID_J00_J99	.0972	.29629
ID_K00_K93	.0871	.28197
ID_L00_L99	.0206	.14205
ID_M00_M99	.0225	.14816
ID_N00_N99	.0713	.25728
ID_O00_O99	.0026	.05128
ID_P00_P96	.0257	.15817
ID_Q00_Q99	.0320	.17607
ID_R00_R99	.0688	.25317
ID_S00_T98	.0862	.28068

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 4.893.

Tabla 135: *Evaluación - Coeficientes del modelo de RLM de Los Ríos*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-95.652	9.807		-9.753	.000	
TYear	.049	.005	.057	9.984	.000	1.005
TMonth	.035	.009	.024	4.141	.000	1.001
DLocationID	-1.157	.061	-.112	-19.079	.000	1.063
SexID	-1.106	.059	-.107	-18.721	.000	1.011
ID_A00_B99	.667	.263	.031	2.538	.011	4.684
ID_C00_D48	10.685	.253	.616	42.173	.000	6.595
ID_E00_E90	1.189	.256	.064	4.641	.000	5.848
ID_F00_F99	.250	.266	.011	.943	.346	4.442
ID_G00_G99	.628	.261	.031	2.410	.016	5.083
ID_I00_I99	12.332	.253	.711	48.677	.000	6.595
ID_J00_J99	3.964	.254	.227	15.634	.000	6.536
ID_K00_K93	2.317	.255	.126	9.068	.000	6.009
ID_L00_L99	-.040	.312	-.001	-.127	.899	2.271
ID_M00_M99	.209	.306	.006	.682	.495	2.387
ID_N00_N99	.631	.260	.031	2.431	.015	5.169
ID_O00_O99	.846	.620	.008	1.365	.172	1.169
ID_P00_P96	.646	.299	.020	2.162	.031	2.586
ID_Q00_Q99	.605	.287	.021	2.106	.035	2.959
ID_R00_R99	1.751	.261	.086	6.698	.000	5.070
ID_S00_T98	4.256	.256	.231	16.629	.000	5.975

Al reemplazar los valores B en la ecuación. el modelo podría definirse de la siguiente manera:

**N°Deaths** = - 95. + 0.49(TYear) + 0.35(TMonth) - 1.157 (DLocationID) - 1.10(SexID) + 0.66(ID\_A00\_B99) + 10.68(ID\_C00\_D48) + 1.18(ID\_E00\_E90) + 0.25(ID\_F00\_F99) + 0.62(ID\_G00\_G99) + 12.33(ID\_I00\_I99) + 3.96(ID\_J00\_J99) + 2.31(ID\_K00\_K93) - 0.04(ID\_L00\_L99) + 0.20(ID\_M00\_M99) + 0.63(ID\_N00\_N99) + 0.84(ID\_O00\_O99) + 0.64(ID\_P00\_P96) + 0.60(ID\_Q00\_Q99) + 1.75(ID\_R00\_R99) + 4.25(ID\_S00\_T98)

Los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.049): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.049 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.035): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.035 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = -1.157): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 1.157, manteniendo las otras condiciones constantes.

**SexID** (B = -1.106): Este valor establece que cuando el género es femenino, la mortalidad cae en 1.106, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.667): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.667, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 10.685): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 10.685, manteniendo constantes las otras condiciones.



**ID\_E00\_E90** (B = 1.189): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 1.189 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_F00\_F99** (B = 0.250): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.250, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.628): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.628 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_I00\_I99** (B = 12.332): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 12.332, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 3.964): Este valor indica que, dado que hay presencia de enfermedades del sistema respiratorio, generará un aumento de la mortalidad de 3,964, manteniendo constantes las demás afecciones.

**ID\_K00\_K93** (B = 2.317): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 2.317, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.209): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.209 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 0.631): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.631, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.846): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.846 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 0.646): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.646, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.605): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.605, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 1.751): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 1.751, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 4.256): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 4.256 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.057$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.03), la mortalidad aumenta en un 5.7% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.294 ( $0.057 * 5.169$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.024$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.43), la mortalidad aumenta en un 2.4% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169,

por lo que esto constituye un cambio de 0.124 ( $0.024 * 5.169$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.112$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 11.2% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.578 ( $0.112 * 5.169$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.107$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 10.7% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.553 ( $0.107 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.031$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.241), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.160 ( $0.031 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.616$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.297), la mortalidad aumenta en un 61.6% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 3.184 ( $0.616 * 5.169$ ), manteniendo constantes las otras condiciones.

**ID\_E00\_E90** ( $\beta = 0.064$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.277), la mortalidad aumenta en un 6.4% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.330 ( $0.064 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.011$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.233), la mortalidad aumenta en un 1.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.056 ( $0.011 * 5.169$ ), manteniendo constantes las otras condiciones.

**ID\_G00\_G99** ( $\beta = 0.031$ ): Este valor establece que, dado que hay enfermedades del sistema nervioso central con una desviación estándar (0.254), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.160 ( $0.031 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.711$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.297), la mortalidad aumenta en un 71.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 3,706 ( $0.717 * 5.169$ ), manteniendo constantes las demás condiciones.

**ID\_J00\_J99** ( $\beta = 0.227$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.296), la mortalidad aumenta en un 22.7% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 1.173 ( $0.227 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.126$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.281), la mortalidad aumenta en un 12.6% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.651 ( $0.126 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.006$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.148), la mortalidad aumenta en un 0.6% de las desviaciones

estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.03 ( $0.006 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.031$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.257), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.160 ( $0.031 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.008$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.051), la mortalidad aumenta en un 0.8% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.041 ( $0.008 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.020$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.158), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.103 ( $0.020 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.021$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.176), la mortalidad aumenta en un 2.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 0.051 ( $0.021 * 5.169$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.086$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.253), la mortalidad aumenta en un 8.6% de las desviaciones estándar. La desviación estándar del número de

muerdes es 5.169, por lo que esto constituye un cambio de 0.444 ( $0.086 * 5.169$ ), manteniendo las otras condiciones constantes

**ID\_S00\_T98** ( $\beta = 0.231$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.280), la mortalidad aumenta en un 23.1% de las desviaciones estándar. La desviación estándar del número de muertes es 5.169, por lo que esto constituye un cambio de 1.194 ( $0.231 * 5.169$ ), manteniendo constantes las otras condiciones.

Tabla 136: *Evaluación - Resumen del modelo de RLM de Atacama*

Modelo	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Atacama	0.534	0.533	2.06940	1.997

Se puede ver que el modelo generado para la región de Atacama presenta un R cuadrado ajustado de 0.533, es decir que la varianza de la mortalidad se explica al 55.3% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 137: *Evaluación - ANOVA del modelo de RLM de Atacama*

Model	Sum of squares	Square Mean	F	Sig.
Regression	41158.393	2057.920	480.549	.000
Residue	35856.815	4.282	-	-
Total	77015.207	-	-	-

La Tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor menor a 0.05.

Tabla 138: *Evaluación - Estadísticos descriptivos del modelo de Atacama*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	3.2686	3.02921
TYear	2007.3307	6.01832
TMonth	6.5606	3.41703
DLocationID	.5412	.49833
SexID	.4797	.49962
ID_A00_B99	.0548	.22761
ID_C00_D48	.1189	.32368
ID_E00_E90	.0811	.27305
ID_F00_F99	.0387	.19293
ID_G00_G99	.0580	.23379
ID_I00_I99	.1195	.32438
ID_J00_J99	.1078	.31017
ID_K00_K93	.0910	.28765
ID_L00_L99	.0093	.09595
ID_M00_M99	.0181	.13335
ID_N00_N99	.0623	.24173
ID_O00_O99	.0018	.04224
ID_P00_P96	.0359	.18595
ID_Q00_Q99	.0351	.18416
ID_R00_R99	.0624	.24194
ID_S00_T98	.0913	.28799

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 3.2686.

Tabla 139: *Evaluación - Coeficientes del modelo de RLM de Atacama*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-97.618	7.569		-12.897	.000	
TYear	.049	.004	.098	13.110	.000	1.008
TMonth	.020	.007	.023	3.079	.002	1.001
DLocationID	-.396	.047	-.065	-8.363	.000	1.089
SexID	-.687	.045	-.113	-15.109	.000	1.012
ID_A00_B99	.357	.214	.027	1.665	.096	4.664
ID_C00_D48	5.525	.202	.590	27.298	.000	8.412
ID_E00_E90	.784	.207	.071	3.785	.000	6.272
ID_F00_F99	.081	.223	.005	.362	.717	3.644
ID_G00_G99	.217	.213	.017	1.015	.310	4.876
ID_I00_I99	5.997	.202	.642	29.641	.000	8.443
ID_J00_J99	2.331	.203	.239	11.452	.000	7.808
ID_K00_K93	1.201	.206	.114	5.845	.000	6.850
ID_L00_L99	-.106	.303	-.003	-.352	.725	1.652
ID_M00_M99	.102	.255	.004	.399	.690	2.260
ID_N00_N99	.390	.212	.031	1.843	.065	5.131
ID_O00_O99	.353	.568	.005	.621	.535	1.128
ID_P00_P96	.554	.226	.034	2.453	.014	3.462
ID_Q00_Q99	.429	.226	.026	1.895	.058	3.406
ID_R00_R99	.556	.212	.044	2.619	.009	5.180
ID_S00_T98	1.992	.206	.189	9.679	.000	6.885

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:



$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 97.618 + 0.04(\text{TYear}) + 0.02(\text{TMonth}) - 0.39(\text{DLocationID}) - 0.68(\text{SexID}) \\ & + 0.35(\text{ID\_A00\_B99}) + 5.52(\text{ID\_C00\_D48}) + 0.78(\text{ID\_E00\_E90}) + 0.08(\text{ID\_F00\_F99}) + \\ & 0.21(\text{ID\_G00\_G99}) + 5.99(\text{ID\_I00\_I99}) + 2.33(\text{ID\_J00\_J99}) + 1.20(\text{ID\_K00\_K93}) - \\ & 0.10(\text{ID\_L00\_L99}) + 0.10(\text{ID\_M00\_M99}) + 0.39(\text{ID\_N00\_N99}) + 0.35(\text{ID\_O00\_O99}) + \\ & 0.55(\text{ID\_P00\_P96}) + 0.42(\text{ID\_Q00\_Q99}) + 0.55(\text{ID\_R00\_R99}) + 1.99(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar hasta qué punto cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.049): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.049 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.020): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.020, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 0.396): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.396, manteniendo las otras condiciones constantes.

**SexID** (B = - 0.687): Este valor establece que cuando el género es femenino, la mortalidad cae en 0.687, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.357): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.357, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 5.525): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 5.525, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 0.784): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 0.784 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.081): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.081, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.217): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.217 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 5.997): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 5.997, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 2.331): Este valor indica que, dado que existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 2.331 en la mortalidad, manteniendo constantes las otras afecciones.

**ID\_K00\_K93** (B = 1.201): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 1.201, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.102): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.102 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 0.390): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.390, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.353): Este valor aclara que como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.353 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 0.554): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.554, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.429): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.429 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 0.556): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.556, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 1.992): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 1.992 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.098$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.01), la mortalidad aumenta en un 9.8% de desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.296 ( $0.098 * 3.029$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.023$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.41), la mortalidad aumenta en un 2.3% de desviaciones

estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.069 ( $0.023 * 3.029$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.067$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 6.7% de las desviaciones estándar. La desviación estándar del número de muertes es de 3.029, por lo que esto constituye un cambio de 0.202 ( $0.067 * 3.029$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.113$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 11.3% de las desviaciones estándar. La desviación estándar del número de muertes es de 3.029, por lo que esto constituye un cambio de 0.342 ( $0.113 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.027$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.227), la mortalidad aumenta en un 2.7% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.081 ( $0.027 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.590$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.323), la mortalidad aumenta en un 59% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 1.787 ( $0.590 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.071$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.273), la mortalidad aumenta en un 7.1% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.215 ( $0.071 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.005$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.192), la mortalidad aumenta en un 0.5% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.015 ( $0.005 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.017$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.233), la mortalidad aumenta en un 1.7% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.051 ( $0.017 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.642$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.324), la mortalidad aumenta en un 64.2% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 1.944 ( $0.642 * 3.029$ ), manteniendo constantes las otras condiciones.

**ID\_J00\_J99** ( $\beta = 0.239$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.310), la mortalidad aumenta en un 23.9% de las desviaciones estándar. La desviación estándar del número de muertes es de 3.029, por lo que esto constituye un cambio de 0.723 ( $0.239 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.114$ ): Este valor indica que cuando hay una presencia de enfermedades del sistema digestivo con una desviación estándar (0.287), la mortalidad aumenta en un 11.4% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.345 ( $0.114 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.004$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.133), la mortalidad aumenta en un 0.4% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.012 ( $0.004 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.031$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.241), la mortalidad aumenta en un 3.1% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.093 ( $0.031 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.005$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.042), la mortalidad aumenta en un 0.5% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.015 ( $0.005 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.034$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.185), la mortalidad aumenta en un 3.4% de las desviaciones estándar. La desviación estándar del número de muertes es de 3.029, por lo que esto constituye un cambio de 0.102 ( $0.034 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.026$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.184), la mortalidad aumenta en un 2.6% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.078 ( $0.026 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.044$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.241), la mortalidad aumenta en un 4.4% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.133 ( $0.044 * 3.029$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.189$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.287), la mortalidad aumenta en un 18.9% de las desviaciones estándar. La desviación estándar del número de muertes es 3.029, por lo que esto constituye un cambio de 0.572 ( $0.189 * 3.029$ ), manteniendo las otras condiciones constantes.

Tabla 140: *Evaluación - Resumen del modelo de RLM de Tarapacá*

<b>Model</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Tarapacá	0.488	0.486	1.97491	2.088

Se puede ver que el modelo generado para la región de Santiago presenta un R cuadrado ajustado de 0.486, es decir que la varianza de mortalidad se explica al 48.6% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 141: *Evaluación - ANOVA del modelo de RLM de Tarapacá*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	30843.243	1542.162	395.399	.000
Residue	32407.357	3.900	-	-
Total	63250.600	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 142: *Evaluación - Estadísticos descriptivos del modelo de Tarapacá*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	2.9517	2.75573
TYear	2007.3810	6.00976
TMonth	6.5264	3.44190
DLocationID	.5255	.49938
SexID	.4881	.49989
ID_A00_B99	.0604	.23821
ID_C00_D48	.1184	.32306
ID_E00_E90	.0802	.27161
ID_F00_F99	.0543	.22655
ID_G00_G99	.0600	.23755
ID_I00_I99	.1206	.32574
ID_J00_J99	.1022	.30288
ID_K00_K93	.0953	.29367
ID_L00_L99	.0104	.10167
ID_M00_M99	.0176	.13167
ID_N00_N99	.0619	.24107
ID_O00_O99	.0025	.05015
ID_P00_P96	.0281	.16524
ID_Q00_Q99	.0318	.17551
ID_R00_R99	.0514	.22079
ID_S00_T98	.0959	.29450

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 2.9517.



Tabla 143: *Evaluación - Coeficientes del modelo de RLM de Tarapacá*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-113.029	7.265		-15.559	.000	
TYear	.057	.004	.124	15.799	.000	1.007
TMonth	.016	.006	.020	2.491	.013	1.001
DLocationID	-.703	.045	-.127	-15.609	.000	1.081
SexID	-.633	.044	-.115	-14.512	.000	1.014
ID_A00_B99	.663	.246	.057	2.697	.007	7.334
ID_C00_D48	5.054	.238	.592	21.206	.000	12.658
ID_E00_E90	.773	.242	.076	3.195	.001	9.232
ID_F00_F99	.477	.248	.039	1.922	.055	6.751
ID_G00_G99	.503	.246	.043	2.043	.041	7.307
ID_I00_I99	5.136	.238	.607	21.567	.000	12.850
ID_J00_J99	1.313	.240	.144	5.482	.000	11.241
ID_K00_K93	1.490	.240	.159	6.203	.000	10.624
ID_L00_L99	.132	.312	.005	.422	.673	2.155
ID_M00_M99	.303	.282	.014	1.077	.282	2.939
ID_N00_N99	.614	.246	.054	2.501	.012	7.486
ID_O00_O99	.458	.489	.008	.936	.349	1.284
ID_P00_P96	.639	.264	.038	2.422	.015	4.059
ID_Q00_Q99	.607	.260	.039	2.335	.020	4.449
ID_R00_R99	.348	.249	.028	1.396	.163	6.477
ID_S00_T98	2.225	.240	.238	9.263	.000	10.692

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ\text{Deaths} = & - 113.029 + 0.057(\text{TYear}) + 0.016 (\text{TMonth}) - 0.70(\text{DLocationID}) - \\ & 0.63(\text{SexID}) + 0.66(\text{ID\_A00\_B99}) + 5.05(\text{ID\_C00\_D48}) + 0.77(\text{ID\_E00\_E90}) + \\ & 0.47(\text{ID\_F00\_F99}) + 0.50(\text{ID\_G00\_G99}) + 5.13(\text{ID\_I00\_I99}) + 1.31(\text{ID\_J00\_J99}) + \\ & 1.49(\text{ID\_K00\_K93}) + 0.13(\text{ID\_L00\_L99}) + 0.30(\text{ID\_M00\_M99}) + 0.61(\text{ID\_N00\_N99}) + \\ & 0.45(\text{ID\_O00\_O99}) + 0.63(\text{ID\_P00\_P96}) + 0.60(\text{ID\_Q00\_Q99}) + 0.34(\text{ID\_R00\_R99}) + \\ & 2.22(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.057): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.057 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.016): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de la mortalidad de 0.016, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 0.703): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.703, manteniendo las otras condiciones constantes.

**SexID** (B = - 0.633): Este valor establece que cuando el género es femenino, la mortalidad cae en 0.633, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.633): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.633, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 5.054): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 5.054, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 0.773): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 0.773 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.477): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.477, manteniendo constantes las otras condiciones.

**ID\_G00\_G99** (B = 0.503): Este valor establece que, dado que existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.503 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 5.136): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 5.136, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 1.313): Este valor indica que, dado que hay presencia en enfermedades del sistema respiratorio, generará un aumento de 1.313 en la mortalidad, manteniendo constantes las otras condiciones

**ID\_K00\_K93** (B = 1.490): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 1.490, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.303): Este valor determina que, como hay presencia en enfermedades del sistema músculo esquelético y del tejido conectivo, generará un aumento de 0.303 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 0.614): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.614, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.458): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.458 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 0.639): Este valor indica que cuando hay presencia en ciertas condiciones que se originan en el período perinatal, la mortalidad permanece en 0.639, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.607): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.607 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_R00\_R99** (B = 0.348): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.348, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 2.225): Este valor establece que como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 2.225 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.124$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.009), la mortalidad aumenta en un 12.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.341 ( $0,124 * 2.755$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.020$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.44), la mortalidad aumenta en un 2% de las desviaciones

estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.05 ( $0.020 * 2.755$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.127$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae en un 12.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.349 ( $0.127 * 2.755$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.115$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en un 11.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.316 ( $0.115 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.057$ ): Este valor aclara que, como hay presencia de ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.238), la mortalidad aumenta en un 5.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.157 ( $0.057 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.592$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.323), la mortalidad aumenta en un 59.2% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 1.630 ( $0.592 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.076$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.271), la mortalidad aumenta en un 7.6% de las desviaciones estándar. La desviación estándar del número de muertes es de 2.755, por lo que esto constituye un cambio de 0.209 ( $0.076 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.039$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.226), la mortalidad aumenta en un 3.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.107 ( $0.039 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.043$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.237), la mortalidad aumenta en un 4.3% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.118 ( $0.043 * 2.755$ ), manteniendo constantes las otras condiciones.

**ID\_I00\_I99** ( $\beta = 0.607$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.325), la mortalidad aumenta en un 60.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 1.672 ( $0.607 * 2.755$ ), manteniendo constantes las demás condiciones.

**ID\_J00\_J99** ( $\beta = 0.144$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.302), la mortalidad aumenta en un 14.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.396 ( $0.144 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.159$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.293), la mortalidad aumenta en un 15.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.438 ( $0.159 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.014$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.293), la mortalidad aumenta en un 15.9% de las desviaciones estándar. La desviación

estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.438 ( $0.159 * 2.755$ ), las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.054$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.241), la mortalidad aumenta en 5.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.148 ( $0.054 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.008$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.050), la mortalidad aumenta en un 0.8% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.022 ( $0.008 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.038$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.165), la mortalidad aumenta en un 3.8% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.104 ( $0.038 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.039$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.175), la mortalidad aumenta en un 3.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.107 ( $0.039 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.028$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.220), la mortalidad aumenta en un 2.8% de las desviaciones estándar. La desviación estándar del número de

muertes es 2.755, por lo que esto constituye un cambio de 0.077 ( $0.028 * 2.755$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.238$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.294), la mortalidad aumenta en un 23.8% de las desviaciones estándar. La desviación estándar del número de muertes es 2.755, por lo que esto constituye un cambio de 0.655 ( $0.238 * 2.755$ ), manteniendo las otras condiciones constantes.

Tabla 144: *Evaluación - Resumen del modelo de RLM de Arica y Parinacota*

Model	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Arica y Parinacota	0.458	0.456	1.83589	2.109

Se puede ver que el modelo generado para la región de Arica y Parinacota presenta un R cuadrado ajustado de 0.458, es decir que la varianza de la mortalidad se explica al 45.6% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón  $H_0$  no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 145: *Evaluación - ANOVA del modelo de RLM de Arica y Parinacota*

Model	Sum of squares	Square Mean	F	Sig.
Regression	22641.837	1132.092	335.884	.000
Residue	26842.538	3.370	-	-
Total	49484.374	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.



Tabla 146: *Evaluación-Estadísticos descriptivos del modelo de Arica y Parinacota*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	2.7451	2.48957
TYear	2007.4912	6.01488
TMonth	6.5082	3.45392
DLocationID	.5043	.50001
SexID	.4784	.49956
ID_A00_B99	.0645	.24565
ID_C00_D48	.1205	.32554
ID_E00_E90	.0907	.28716
ID_F00_F99	.0377	.19047
ID_G00_G99	.0549	.22771
ID_I00_I99	.1251	.33086
ID_J00_J99	.1056	.30731
ID_K00_K93	.0991	.29876
ID_L00_L99	.0073	.08492
ID_M00_M99	.0200	.14014
ID_N00_N99	.0678	.25134
ID_O00_O99	.0030	.05474
ID_P00_P96	.0220	.14683
ID_Q00_Q99	.0278	.16442
ID_R00_R99	.0471	.21184
ID_S00_T98	.0956	.29400

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 2.7451.

Tabla 147: *Evaluación - Coeficientes del modelo de RLM de Arica y Parinacota*

<b>Variable</b>	<b>B</b>	<b>Desv. Error</b>	<b><math>\beta</math></b>	<b>T</b>	<b>Sig.</b>	<b>VIF</b>
(Constant)	-85.278	6.883		-12.390	.000	
TYear	.043	.003	.105	12.660	.000	1.007
TMonth	.005	.006	.007	.815	.415	1.001
DLocationID	-.832	.043	-.167	-19.444	.000	1.085
SexID	-.465	.041	-.093	-11.222	.000	1.016
ID_A00_B99	.355	.208	.035	1.707	.088	6.178
ID_C00_D48	4.653	.201	.608	23.204	.000	10.096
ID_E00_E90	.685	.203	.079	3.367	.001	8.077
ID_F00_F99	-.117	.220	-.009	-.531	.595	4.141
ID_G00_G99	.188	.211	.017	.890	.373	5.461
ID_I00_I99	3.763	.200	.500	18.799	.000	10.391
ID_J00_J99	1.037	.202	.128	5.141	.000	9.107
ID_K00_K93	1.110	.202	.133	5.487	.000	8.655
ID_L00_L99	-.231	.308	-.008	-.750	.453	1.620
ID_M00_M99	.000	.240	.000	.001	.999	2.689
ID_N00_N99	.245	.207	.025	1.185	.236	6.421
ID_O00_O99	.445	.422	.010	1.057	.291	1.262
ID_P00_P96	.424	.237	.025	1.793	.073	2.859
ID_Q00_Q99	.380	.228	.025	1.667	.096	3.323
ID_R00_R99	.076	.215	.006	.354	.723	4.899
ID_S00_T98	1.474	.203	.174	7.264	.000	8.432

Al reemplazar los valores B en la ecuación. el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 85.278 + 0.043(\text{TYear}) + 0.005(\text{TMonth}) - 0.83(\text{DLocationID}) - 0.46(\text{SexID}) \\ & + 0.35(\text{ID\_A00\_B99}) + 4.65(\text{ID\_C00\_D48}) + 0.68(\text{ID\_E00\_E90}) - 0.11(\text{ID\_F00\_F99}) + \\ & 0.188(\text{ID\_G00\_G99}) + 3.76(\text{ID\_I00\_I99}) + 1.03(\text{ID\_J00\_J99}) + 1.11(\text{ID\_K00\_K93}) - \\ & 0.23(\text{ID\_L00\_L99}) + 0.24(\text{ID\_N00\_N99}) + 0.44(\text{ID\_O00\_O99}) + 0.42(\text{ID\_P00\_P96}) + \\ & 0.38(\text{ID\_Q00\_Q99}) + 0.07(\text{ID\_R00\_R99}) + 1.474(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.043): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.043 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.005): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de mortalidad de 0.005, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 0.832): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.832, manteniendo las otras condiciones constantes.

**SexID** (B = - 0.465): Este valor establece que cuando el género es femenino, la mortalidad cae en 0.465, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.355): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.355, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 4.653): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 4.653, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 0.687): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 0.687 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.188): Este valor establece que, dado que existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.188 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 3.763): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 3.763, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 1.037): Este valor indica que, dado que hay una presencia en enfermedades del sistema respiratorio, generará un aumento de 1.037 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_K00\_K93** (B = 1.11): Este valor indica que, dado que hay una presencia en enfermedades del sistema respiratorio, generar un aumento de 1.11 en la mortalidad, transmitir constantes las otras condiciones.

**ID\_N00\_N99** (B = 0.245): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.245, manteniendo constantes las otras condiciones.

**ID\_O00\_O99** (B = 0.445): Este valor aclara que, dado que existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.445 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_P00\_P96** (B = 0.424): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.424, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.380): Este valor indica que debido a la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará

un aumento de 0.380 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $B = 0.076$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anomalías clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.076, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $B = 1.474$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 1.474 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.105$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.01), la mortalidad aumenta en un 10.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.261 ( $0.105 * 2.489$ ), manteniendo las otras condiciones constantes,

**TMonth** ( $\beta = 0.007$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.45), la mortalidad aumenta en un 0.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.017 ( $0.007 * 2.489$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.167$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.50), la mortalidad cae un 16.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.415 ( $0.167 * 2.489$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.093$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae un 9.3% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.231 ( $0.093 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.035$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.245), la mortalidad aumenta en un 3.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.087 ( $0.035 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.608$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.325), la mortalidad aumenta en un 60.8% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 1.513 ( $0.608 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.079$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.287), la mortalidad aumenta en un 7.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.196 ( $0.079 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.017$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.227), la mortalidad aumenta en un 1.7% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.042 ( $0.017 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.50$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.330), la mortalidad aumenta en un 50% de las desviaciones estándar. La desviación

estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 1.244 ( $0.50 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.128$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.307), la mortalidad aumenta en un 12.8% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.318 ( $0.128 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.133$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.298), la mortalidad aumenta en un 13.3% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.331 ( $0.133 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.025$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.251), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.062 ( $0.025 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.010$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.054), la mortalidad aumenta en un 0.1% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.024 ( $0.010 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.025$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.146), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.062 ( $0.025 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.025$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.164), la mortalidad aumenta en un 2.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.014 ( $0.006 * 2.489$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.006$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.211), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.014 ( $0.006 * 2.489$ ), manteniendo las otras condiciones constantes

**ID\_S00\_T98** ( $\beta = 0.174$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.294), la mortalidad aumenta en un 17.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.489, por lo que esto constituye un cambio de 0.433 ( $0.174 * 2.489$ ), manteniendo las otras condiciones constantes.

Tabla 148: *Evaluación - Resumen del modelo de RLM de Magallanes*

Model	R Square	Adjusted R Square	Standard error of the estimate	Durbin - Watson
Magallanes	0.531	0.530	1.71180	2.035

Se puede ver que el modelo generado para la región de Magallanes y de la Antártica Chilena presenta un R cuadrado ajustado de 0.530, es decir que la varianza de mortalidad se explica al 53% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.



Tabla 149: *Evaluación - ANOVA del modelo de RLM de Magallanes*

<b>Model</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	23737.469	1186.873	405.042	.000
Residue	20942.483	2.930	-	-
Total	44679.952	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 150: *Evaluación - Estadísticos descriptivos del modelo de Magallanes*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	2.8520	2.49682
TYear	2007.2023	6.07507
TMonth	6.5167	3.41183
DLocationID	.5612	.49627
SexID	.4834	.49976
ID_A00_B99	.0453	.20806
ID_C00_D48	.1388	.34577
ID_E00_E90	.0774	.26729
ID_F00_F99	.0558	.22956
ID_G00_G99	.0571	.23197
ID_I00_I99	.1402	.34723
ID_J00_J99	.1088	.31143
ID_K00_K93	.1024	.30319
ID_L00_L99	.0114	.10635
ID_M00_M99	.0153	.12293
ID_N00_N99	.0601	.23774
ID_O00_O99	.0011	.03339
ID_P00_P96	.0201	.14032
ID_Q00_Q99	.0236	.15174
ID_R00_R99	.0333	.17954
ID_S00_T98	.0964	.29516

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 2.8520.

Tabla 151: *Evaluación - Coeficientes del modelo de RLM de Magallanes*

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-35.486	6.709		-5.289	.000	
TYear	.018	.003	.044	5.471	.000	1.006
TMonth	.021	.006	.029	3.545	.000	1.001
DLocationID	-.014	.042	-.003	-.324	.746	1.069
SexID	-.657	.041	-.132	-16.124	.000	1.015
ID_A00_B99	.354	.203	.029	1.741	.082	4.368
ID_C00_D48	4.180	.188	.579	22.266	.000	10.307
ID_E00_E90	.669	.194	.072	3.455	.001	6.556
ID_F00_F99	.357	.199	.033	1.794	.073	5.112
ID_G00_G99	.261	.199	.024	1.314	.189	5.192
ID_I00_I99	5.051	.188	.702	26.918	.000	10.384
ID_J00_J99	1.272	.190	.159	6.699	.000	8.547
ID_K00_K93	1.240	.190	.151	6.515	.000	8.143
ID_L00_L99	.245	.261	.010	.940	.347	1.883
ID_M00_M99	.229	.243	.011	.942	.346	2.179
ID_N00_N99	.466	.198	.044	2.361	.018	5.393
ID_O00_O99	.337	.632	.005	.534	.593	1.088
ID_P00_P96	.232	.230	.013	1.009	.313	2.539
ID_Q00_Q99	.104	.223	.006	.466	.641	2.793
ID_R00_R99	.197	.212	.014	.932	.351	3.535
ID_S00_T98	1.610	.191	.190	8.416	.000	7.794

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 35.486 + 0.01(\text{TYear}) + 0.02(\text{TMonth}) - 0.01 (\text{DLocationID}) - 0.65(\text{SexID}) \\ & + 0.35(\text{ID\_A00\_B99}) + 4.18(\text{ID\_C00\_D48}) + 0.66(\text{ID\_E00\_E90}) + 0.35(\text{ID\_F00\_F99}) + \\ & 0.26(\text{ID\_G00\_G99}) + 5.05(\text{ID\_I00\_I99}) + 1.27(\text{ID\_J00\_J99}) + 1.24(\text{ID\_K00\_K93}) + \\ & 0.24(\text{ID\_L00\_L99}) + 0.22(\text{ID\_M00\_M99}) + 0.46(\text{ID\_N00\_N99}) + 0.33(\text{ID\_O00\_O99}) + \\ & 0.23(\text{ID\_P00\_P96}) + 0.10(\text{ID\_Q00\_Q99}) + 0.19(\text{ID\_R00\_R99}) + 1.61(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.018): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.018 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.021): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de 0.021 en la mortalidad, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 0.014): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.014, manteniendo las otras condiciones constantes.

**SexID** (B = - 0.657): Este valor establece que cuando el género es femenino, la mortalidad cae en 0.657, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.354): Este valor aclara que ante la presencia de determinadas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.354, manteniendo constantes las demás condiciones.

**ID\_C00\_D48** (B= 4.180): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 4.180, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 0.669): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 0.669 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_F00\_F99** (B = 0.357): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento, la mortalidad aumenta a 0.357, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.261): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.261 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 5.051): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 5.051, manteniendo las otras condiciones constantes.

**ID\_J00\_J99** (B = 1.272): Este valor indica que, como existe una presencia en enfermedades del sistema respiratorio, generará un aumento de 1.272 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_K00\_K93** (B = 1.240): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 1.240, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.229): Este valor determina que, como existe una presencia en enfermedades del sistema musculoesquelético y del tejido conectivo, generará un aumento de 0.229 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_N00\_N99** (B = 0.466): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.466, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.337): Este valor aclara que como existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.337 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 0.232): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.232, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.104): Este valor indica que, como existe una presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.104 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_R00\_R99** (B = 0.197): Este valor determina que cuando existe la presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no clasificada en otra parte, la mortalidad permanece en 0.197, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** (B = 1.610): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 1.610 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.044$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (6.07), la mortalidad aumenta en un 4.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.109 ( $0.044 * 2.496$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.029$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.41), la mortalidad aumenta en un 2.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.072 ( $0.029 * 2.496$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.003$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.496), la mortalidad cae en un 0.3% de las desviaciones estándar. La desviación estándar del número de

mueres es 2.496, por lo que esto constituye un cambio de 0.007 ( $0.003 * 2.496$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.132$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.50), la mortalidad cae en 13.2% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.329 ( $0.132 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.029$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.208), la mortalidad aumenta en un 2.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.072 ( $0.029 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.579$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.345), la mortalidad aumenta en un 57.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 1.445 ( $0.579 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.072$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.267), la mortalidad aumenta en un 7.2% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.179 ( $0.072 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_F00\_F99** ( $\beta = 0.033$ ): Este valor determina que cuando hay presencia de enfermedades de trastornos mentales y del comportamiento con una desviación estándar (0.229), la mortalidad aumenta en un 3.3% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.082 ( $0.033 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.024$ ): Este valor establece que, dado que hay enfermedades del sistema nervioso central con una desviación estándar (0.231), la mortalidad aumenta en un 2.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.059 ( $0.024 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.702$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.347), la mortalidad aumenta en un 70.2% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 1.752 ( $0.702 * 2.496$ ), manteniendo las otras condiciones constantes

**ID\_J00\_J99** ( $\beta = 0.159$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.311), la mortalidad aumenta en un 15.9% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 1.295 ( $0.159 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.151$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.303), la mortalidad aumenta en un 15.1% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.376 ( $0.151 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.011$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.122), la mortalidad aumenta en un 1.1% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.027 ( $0.011 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.044$ ): Este valor establece que cuando hay presencia de enfermedades del sistema genitourinario con una desviación estándar (0.237),

la mortalidad aumenta en un 4.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.109 ( $0.044 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.005$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.033), la mortalidad aumenta en un 0.5% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.012 ( $0.005 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.013$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.140), la mortalidad aumenta en un 1.3% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.032 ( $0.013 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.006$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.151), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.014 ( $0.006 * 2.496$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.014$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.179), la mortalidad aumenta en un 1.4% de las desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.034 ( $0.014 * 2.496$ ), manteniendo las otras condiciones constantes

**ID\_S00\_T98** ( $\beta = 0.190$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.295), la mortalidad aumenta en un 19% las



desviaciones estándar. La desviación estándar del número de muertes es 2.496, por lo que esto constituye un cambio de 0.474 ( $0.190 * 2.496$ ), manteniendo las otras condiciones constantes.

Tabla 152: *Evaluación - Resumen del modelo de RLM de Aysen*

<b>Model</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Standard error of the estimate</b>	<b>Durbin - Watson</b>
Aysen model	0.280	0.277	1.24746	1.994

Se puede ver que el modelo generado para la región de Aysen presenta una R al cuadrado ajustada de 0.277, es decir que la varianza de la mortalidad se explica al 27.7% por las variables independientes. Por otro lado, en la prueba del estadístico de Durbin Watson, el valor está dentro del rango de 1.5 a 2.5, por esta razón H0 no se rechaza, lo que indica que los residuos no están autocorrelacionados.

Tabla 153: *Evaluación - ANOVA del modelo de RLM de Aysen*

<b>Modelo</b>	<b>Sum of squares</b>	<b>Square Mean</b>	<b>F</b>	<b>Sig.</b>
Regression	3042.029	152.101	97.741	.000
Residue	7813.483	1.556	-	-
Total	10855.	-	-	-

La tabla ANOVA indica que el modelo tiene un ajuste significativo porque presenta un valor inferior a 0.05.

Tabla 154: *Evaluación - Estadísticos descriptivos del modelo de Aysen*

<b>Variable</b>	<b>Mean</b>	<b>Standard Deviation</b>
Quantity	1.9778	1.46746
TYear	2007.3195	5.98796
TMonth	6.6061	3.40378
DLocationID	.5438	.49812

Variable	Mean	Standard Deviation
SexID	.4381	.49621
ID_A00_B99	.0359	.18606
ID_C00_D48	.1787	.38314
ID_E00_E90	.0526	.22317
ID_F00_F99	.0329	.17845
ID_G00_G99	.0385	.19236
ID_I00_I99	.1720	.37738
ID_J00_J99	.1204	.32545
ID_K00_K93	.0849	.27874
ID_L00_L99	.0081	.08982
ID_M00_M99	.0127	.11196
ID_N00_N99	.0430	.20296
ID_O00_O99	.0020	.04449
ID_P00_P96	.0220	.14675
ID_Q00_Q99	.0212	.14414
ID_R00_R99	.0502	.21833
ID_S00_T98	.1168	.32124

En esta tabla puede ver la media y la desviación estándar de cada variable en el conjunto de datos, por lo que se puede ver que el número promedio de muertes es 1.9778.

Tabla 155: *Evaluación - Coeficientes del modelo de RLM de Aysen*

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
(Constant)	-28.663	5.923		-4.839	.000	
TYear	.015	.003	.061	5.089	.000	1.010
TMonth	.001	.005	.002	.184	.854	1.002
DLocationID	-.314	.037	-.107	-8.465	.000	1.108
SexID	-.497	.036	-.168	-13.861	.000	1.025
ID_A00_B99	.155	.216	.020	.721	.471	5.222

Región	B	Desv. Error	$\beta$	T	Sig.	VIF
ID_C00_D48	1.732	.199	.452	8.682	.000	18.920
ID_E00_E90	.176	.209	.027	.842	.400	7.077
ID_F00_F99	-.008	.218	-.001	-.036	.971	4.893
ID_G00_G99	.032	.215	.004	.148	.883	5.530
ID_I00_I99	1.683	.200	.433	8.433	.000	18.379
ID_J00_J99	.602	.201	.133	2.988	.003	13.918
ID_K00_K93	.496	.204	.094	2.431	.015	10.472
ID_L00_L99	-.056	.276	-.003	-.203	.839	1.987
ID_M00_M99	.077	.250	.006	.310	.756	2.533
ID_N00_N99	.135	.212	.019	.634	.526	6.023
ID_O00_O99	.321	.440	.010	.728	.467	1.244
ID_P00_P96	.243	.228	.024	1.063	.288	3.632
ID_Q00_Q99	.097	.229	.010	.423	.672	3.535
ID_R00_R99	.089	.211	.013	.424	.672	6.883
ID_S00_T98	1.295	.202	.284	6.411	.000	13.647

Al reemplazar los valores B en la ecuación, el modelo podría definirse de la siguiente manera:

$$\begin{aligned} \text{N}^\circ \text{Deaths} = & - 28.663 + 0.015(\text{TYear}) + 0.001(\text{TMonth}) - 0.314(\text{DLocationID}) - \\ & 0.497(\text{SexID}) + 0.155(\text{ID\_A00\_B99}) + 1.732(\text{ID\_C00\_D48}) + 0.17(\text{ID\_E00\_E90}) - \\ & 0.008(\text{ID\_F00\_F99}) + 0.032(\text{ID\_G00\_G99}) + 1.683(\text{ID\_I00\_I99}) + 0.602(\text{ID\_J00\_J99}) + \\ & 0.496(\text{ID\_K00\_K93}) - 0.056(\text{ID\_L00\_L99}) + 0.077(\text{ID\_M00\_M99}) + 0.135(\text{ID\_N00\_N99}) \\ & + 0.32(\text{ID\_O00\_O99}) + 0.24(\text{ID\_P00\_P96}) + 0.097(\text{ID\_Q00\_Q99}) + 0.089(\text{ID\_R00\_R99}) \\ & + 1.295(\text{ID\_S00\_T98}) \end{aligned}$$

Los valores B pueden indicar en qué medida cada predictor puede afectar el resultado, por lo que estos valores pueden interpretarse de la siguiente manera:

**Tyear** (B = 0.015): Este valor indica que a medida que el año aumenta en una unidad, la mortalidad aumenta en 0.015 manteniendo las otras condiciones constantes.

**TMonth** (B = 0.001): Este valor indica que a medida que el mes aumenta en una unidad, genera un crecimiento de mortalidad de 0.001, manteniendo las otras condiciones constantes.

**DLocationID** (B = - 0.314): Este valor determina que cuando hay presencia en un hospital o clínica, la mortalidad cae en 0.314, manteniendo las otras condiciones constantes.

**SexID** (B = - 0.497): Este valor establece que cuando el género es femenino, la mortalidad cae en 0.497, manteniendo las otras condiciones constantes.

**ID\_A00\_B99** (B = 0.155): Este valor aclara que, como existe una presencia en ciertas enfermedades infecciosas y parasitarias, la mortalidad aumenta en 0.155, manteniendo constantes las otras condiciones.

**ID\_C00\_D48** (B= 1.732): Este valor indica que cuando hay presencia en enfermedades tumorales [neoplasias], la mortalidad aumenta a 1.732, manteniendo las otras condiciones constantes.

**ID\_E00\_E90** (B = 0.176): Este valor indica que, como existe una presencia en enfermedades endocrinas, nutricionales y metabólicas, generará un aumento de 0.176 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_G00\_G99** (B = 0.032): Este valor establece que, como existe una presencia en enfermedades del sistema nervioso central, generará un aumento de 0.032 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_I00\_I99** (B = 1.683): Este valor aclara que cuando hay presencia en enfermedades del sistema circulatorio, la mortalidad aumenta a 1.683, manteniendo constantes las otras condiciones.

**ID\_J00\_J99** (B = 0.602): Este valor indica que, dado que hay presencia de enfermedades del sistema respiratorio, generará un aumento de 0.602 en la mortalidad, manteniendo constantes las demás afecciones.

**ID\_K00\_K93** (B = 0.496): Este valor indica que cuando hay presencia en enfermedades del sistema digestivo, la mortalidad aumenta a 0.496, manteniendo las otras condiciones constantes.

**ID\_M00\_M99** (B = 0.077): Este valor determina que, como existe una presencia en enfermedades del sistema musculo esquelético y del tejido conectivo, generará un aumento de 0.077 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_N00\_N99** (B = 0.135): Este valor establece que cuando hay presencia en enfermedades del sistema genitourinario, la mortalidad aumenta a 0.135, manteniendo las otras condiciones constantes.

**ID\_O00\_O99** (B = 0.321): Este valor aclara que, dado que existe la presencia de enfermedades del embarazo, el parto y el puerperio, generará un aumento de 0.321 en la mortalidad, manteniendo las otras condiciones constantes.

**ID\_P00\_P96** (B = 0.243): Este valor indica que cuando hay presencia en ciertas condiciones originadas en el período perinatal, la mortalidad permanece en 0.243, manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** (B= 0.097): Este valor indica que debido a la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas, generará un aumento de 0.097 en la mortalidad, manteniendo constantes las otras condiciones.

**ID\_R00\_R99** (B = 0.089): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos de anormalidades clínicas y de laboratorio, no

clasificadas en otra parte, la mortalidad permanece en 0.089, manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $B = 1.295$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas, generará un aumento de 1.295 en la mortalidad, manteniendo las otras condiciones constantes.

Para comprender mejor la importancia de cada variable, el valor  $\beta$  se interpretó de la siguiente manera:

**Tyear** ( $\beta = 0.061$ ): Este valor indica que a medida que el año aumenta en una desviación estándar (5.987), la mortalidad aumenta en un 6.1% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.089 ( $0.061 * 1.467$ ), manteniendo las otras condiciones constantes.

**TMonth** ( $\beta = 0.002$ ): Este valor indica que cuando el mes aumenta en una desviación estándar (3.403), la mortalidad aumenta en un 0.2% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.002 ( $0.002 * 1.467$ ), manteniendo las otras condiciones constantes.

**DLocationID** ( $\beta = - 0.107$ ): Este valor determina que, como existe una presencia en un hospital o clínica con una desviación estándar (0.498), la mortalidad cae en un 10.7% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.156 ( $0.107 * 1.467$ ), manteniendo las otras condiciones constantes.

**SexID** ( $\beta = - 0.168$ ): Este valor establece que cuando hay presencia del género femenino con una desviación estándar (0.496), la mortalidad cae en un 16.8% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.246 ( $0.168 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_A00\_B99** ( $\beta = 0.020$ ): Este valor aclara que, dado que hay ciertas enfermedades infecciosas y parasitarias con una desviación estándar (0.186), la mortalidad aumenta en un 2% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.029 ( $0.020 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_C00\_D48** ( $\beta = 0.452$ ): Este valor indica que cuando hay enfermedades tumorales [neoplasias] con una desviación estándar (0.383), la mortalidad aumenta en un 45.2% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.663 ( $0.452 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_E00\_E90** ( $\beta = 0.027$ ): Este valor indica que, como existe la presencia de enfermedades endocrinas, nutricionales y metabólicas con una desviación estándar (0.223), la mortalidad aumenta en un 2.7% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.039 ( $0.027 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_G00\_G99** ( $\beta = 0.004$ ): Este valor establece que, como hay presencia de enfermedades del sistema nervioso central con una desviación estándar (0.192), la mortalidad aumenta en un 0.4% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.005 ( $0.004 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_I00\_I99** ( $\beta = 0.433$ ): Este valor aclara que cuando hay presencia de enfermedades del sistema circulatorio con una desviación estándar (0.377), la mortalidad aumenta en un 43.3% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.649 ( $0.443 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_J00\_J99** ( $\beta = 0.133$ ): Este valor indica que, dado que hay enfermedades del sistema respiratorio con una desviación estándar (0.325), la mortalidad aumenta

en un 13.3% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.195 ( $0.133 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_K00\_K93** ( $\beta = 0.094$ ): Este valor indica que cuando hay presencia de enfermedades del sistema digestivo con una desviación estándar (0.278), la mortalidad aumenta en un 9.4% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.137 ( $0.094 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_M00\_M99** ( $\beta = 0.006$ ): Este valor determina que, dado que hay enfermedades del sistema musculoesquelético y del tejido conectivo con una desviación estándar (0.111), la mortalidad aumenta en un 0.6% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.008 ( $0.006 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_N00\_N99** ( $\beta = 0.019$ ): Este valor establece que cuando hay una presencia de enfermedades del sistema genitourinario con una desviación estándar (0.202), la mortalidad aumenta en un 1.9% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.027 ( $0.019 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_O00\_O99** ( $\beta = 0.010$ ): Este valor aclara que, como existe la presencia de enfermedades del embarazo, el parto y el puerperio con una desviación estándar (0.044), la mortalidad aumenta en un 1% las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.014 ( $0.010 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_P00\_P96** ( $\beta = 0.024$ ): Este valor indica que cuando hay ciertas condiciones que se originan en el período perinatal con una desviación estándar (0.146), la mortalidad aumenta en un 2.4% de las desviaciones estándar. La desviación



estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.035 ( $0.024 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_Q00\_Q99** ( $\beta = 0.010$ ): Este valor indica que, como existe la presencia de malformaciones congénitas, deformidades y anomalías cromosómicas con una desviación estándar (0.144), la mortalidad aumenta en un 1% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.014 ( $0.010 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_R00\_R99** ( $\beta = 0.013$ ): Este valor determina que cuando hay presencia de síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte con una desviación estándar (0.218), la mortalidad aumenta en un 1.3% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.019 ( $0.013 * 1.467$ ), manteniendo las otras condiciones constantes.

**ID\_S00\_T98** ( $\beta = 0.284$ ): Este valor establece que, como existe la presencia de lesiones, intoxicaciones y algunas otras consecuencias de causas externas con una desviación estándar (0.321), la mortalidad aumenta en un 28.4% de las desviaciones estándar. La desviación estándar del número de muertes es 1.467, por lo que esto constituye un cambio de 0.416 ( $0.284 * 1.467$ ), manteniendo las otras condiciones constantes.

Tabla 156: *Evaluación - Resumen de los modelos de RLM por cada región*

<b>Región</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Durbin- Watson</b>
Valparaíso	0.822	0.822	1.943
Metropolitana de Santiago	0.814	0.813	1.896
Biobío	0.787	0.787	1.903
Maule	0.748	0.748	1.944
Libertador General Bernardo O'Higgins	0.722	0.721	1.943
La Araucanía	0.702	0.701	1.871
Coquimbo	0.699	0.698	2.089
Ñuble	0.690	0.689	2.015
Antofagasta	0.679	0.678	2.022
Los Lagos	0.678	0.677	2.027
Los Ríos	0.670	0.669	1.967
Atacama	0.534	0.533	1.997
Magallanes y de la Antártica Chilena	0.531	0.530	2.035
Tarapacá	0.488	0.486	2.088
Arica y Parinacota	0.458	0.456	2.109
Aysén del Gral. C. Ibañez del Campo	0.280	0.277	1.944

### **Regresión logística ordinal**

La segunda técnica utilizada para explicar las muertes es la regresión logística ordinal, luego se mostrarán los resultados de cada región de Chile, que mostrará la tabla resumen del modelo con el pseudo r.

**Región: Santiago**

Tabla 157: *Evaluación - Coeficientes del modelo de RLO de Santiago*

<b>Variab</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-178.0611	10.9764	-16.22	263.09	<0.0001
<b>y&gt;=3</b>	-180.7034	10.9867	-16.45	270.60	<0.0001
<b>y&gt;=4</b>	-183.6297	11.0010	-16.69	278.56	<0.0001
<b>y&gt;=5</b>	-185.5128	11.0070	-16.85	283.92	<0.0001
<b>TYear</b>	0.0881	0.0055	16.12	259.85	<0.0001
<b>TMonth</b>	0.0254	0.0092	2.78	7.73	0.0055
<b>DLocationID</b>	-2.3214	0.0790	-29.38	863.18	0.0055
<b>SexID</b>	0.1425	0.0639	2.23	4.97	0.0257
<b>RAgeID_0_10</b>	-16.4088	57.7516	-0.28	0.08	0.7763
<b>RAgeID_11_19</b>	-17.2286	75.6517	-0.23	0.05	0.8199
<b>RAgeID_20_64</b>	-4.0368	0.1013	-39.87	1589.62	<0.0001
<b>ID_A00_B99</b>	-14.5276	95.7256	-0.15	0.02	0.8794
<b>ID_C00_D48</b>	3.0847	0.1189	25.94	672.88	<0.0001
<b>ID_D50_D89</b>	-14.9272	130.0819	-0.11	0.01	0.9086
<b>ID_E00_E90</b>	-4.7171	0.7147	-6.60	43.56	<0.0001
<b>ID_F00_F99</b>	-5.4017	1.0054	-5.37	28.84	<0.0001
<b>ID_G00_G99</b>	-14.2657	83.6298	-0.17	0.03	0.8646
<b>ID_I00_I99</b>	4.1700	0.1243	33.54	1124.93	<0.0001
<b>ID_J00_J99</b>	0.1041	0.1293	0.80	0.64	0.4209
<b>ID_K00_K93</b>	-14.6026	98.9180	-0.15	0.02	0.8826
<b>ID_L00_L99</b>	-15.4166	171.5767	-0.09	0.01	0.9284
<b>ID_M00_M99</b>	-14.8696	124.1828	-0.12	0.01	0.9047
<b>ID_N00_N99</b>	-14.8014	109.5404	-0.14	0.02	0.8925
<b>ID_O00_O99</b>	-11.2451	345.5779	-0.03	0.00	0.9740
<b>ID_P00_P96</b>	-7.8732	169.6528	-0.05	0.00	0.9630
<b>ID_Q00_Q99</b>	-12.8822	103.7241	-0.12	0.01	0.9012
<b>ID_R00_R99</b>	-14.4996	94.7654	-0.15	0.02	0.8784

Tabla 158: *Evaluación - Probabilidades del modelo de RLO de Santiago*

<b>Variab</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-1.780.611	-
<b>y&gt;=3</b>	-1.807.034	-
<b>y&gt;=4</b>	-1.836.297	-
<b>y&gt;=5</b>	-1.855.128	-
<b>TYear</b>	0.0881	1.092097326
<b>TMonth</b>	0.0254	1.025725329
<b>DLocationID</b>	-2.321	0.098136099
<b>SexID</b>	0.1425	1.153153081
<b>RAgeID_0_10</b>	-16.4088	7.47737E-08
<b>RAgeID_11_19</b>	-17.2286	3.29393E-08
<b>RAgeID_20_64</b>	-4.0368	0.017653875
<b>ID_A00_B99</b>	-14.5276	4.90618E-07
<b>ID_C00_D48</b>	3.0847	21.86090759
<b>ID_D50_D89</b>	-14.9272	3.29003E-07
<b>ID_E00_E90</b>	-4.7171	0.00894107
<b>ID_F00_F99</b>	-5.4017	0.004508909
<b>ID_G00_G99</b>	-14.2657	6.37507E-07
<b>ID_I00_I99</b>	4.17	64.71545211
<b>ID_J00_J99</b>	0.1041	1.109711421
<b>ID_K00_K93</b>	-14.6026	4.55168E-07
<b>ID_L00_L99</b>	-15.4166	2.01677E-07
<b>ID_M00_M99</b>	-14.8696	3.4851E-07
<b>ID_N00_N99</b>	-14.8014	3.73107E-07
<b>ID_O00_O99</b>	-11.2451	1.30712E-05
<b>ID_P00_P96</b>	-78.732	6.41397E-35
<b>ID_Q00_Q99</b>	-12.8822	2.54291E-06
<b>ID_R00_R99</b>	-14.4996	5.04549E-07

## Región: Valparaíso

Tabla 159: Evaluación - Coeficientes del modelo de RLO de Valparaíso

Variables	Coef	S.E	Wald Z	Wald	Pr(> z )
<b>y&gt;=2</b>	-162.0594	11.4281	-14.18	201.07	<0.0001
<b>y&gt;=3</b>	-164.2689	11.4370	-14.36	206.21	<0.0001
<b>y&gt;=4</b>	-167.2142	11.4504	-14.60	213.16	<0.0001
<b>y&gt;=5</b>	-172.0126	11.4967	-14.96	223.80	<0.0001
<b>TYear</b>	0.0796	0.0057	13.99	195.72	<0.0001
<b>TMonth</b>	0.0394	0.0096	4.12	16.97	<0.0001
<b>DLocationID</b>	-1.6696	0.0749	-22.28	496.40	<0.0001
<b>SexID</b>	0.0107	0.0665	0.16	0.03	0.8718
<b>RAgeID_0_10</b>	-14.8020	46.5892	-0.32	0.10	0.7507
<b>RAgeID_11_19</b>	-15.8191	66.3770	-0.24	0.06	0.8116
<b>RAgeID_20_64</b>	-4.5121	0.1109	-40.70	1656.49	<0.0001
<b>ID_A00_B99</b>	-12.4566	61.6745	-0.20	0.04	0.8399
<b>ID_C00_D48</b>	3.9337	0.1450	27.14	736.58	<0.0001
<b>ID_D50_D89</b>	-12.8522	97.7578	-0.13	0.02	0.8954
<b>ID_E00_E90</b>	-4.4350	1.0087	-4.40	19.36	<0.0001
<b>ID_F00_F99</b>	-13.1575	79.0866	-0.17	0.03	0.8679
<b>ID_G00_G99</b>	-12.4642	56.0812	-0.22	0.05	0.8241
<b>ID_I00_I99</b>	5.1980	0.1527	34.04	1158.72	<0.0001
<b>ID_J00_J99</b>	0.8725	0.1548	5.63	31.70	<0.0001
<b>ID_K00_K93</b>	-3.0438	0.5171	-5.89	34.69	<0.0001
<b>ID_L00_L99</b>	-13.3223	114.7066	-0.12	0.01	0.9075
<b>ID_M00_M99</b>	-12.7434	92.9380	-0.14	0.02	0.8909
<b>ID_N00_N99</b>	-12.8375	66.9934	-0.19	0.04	0.8480
<b>ID_O00_O99</b>	-8.5912	303.1920	-0.03	0.001	0.9774
<b>ID_P00_P96</b>	-3.3835	105.2273	-0.05	0.003	0.9743
<b>ID_Q00_Q99</b>	-10.1618	77.4986	-0.13	0.02	0.8957
<b>ID_R00_R99</b>	-12.7384	66.0539	-0.19	0.04	0.8471

Tabla 160: *Evaluación - Probabilidades del modelo de RLO de Valparaíso*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-162.0594	-
<b>y&gt;=3</b>	-164.269	-
<b>y&gt;=4</b>	-167.214	-
<b>y&gt;=5</b>	-172.0126	-
<b>TYear</b>	0.0796	1.082853839
<b>TMonth</b>	0.0394	1.040186475
<b>DLocationID</b>	-1.6696	0.18832238
<b>SexID</b>	0.0107	1.01075745
<b>RAgeID_0_10</b>	-14.802	3.72883E-07
<b>RAgeID_11_19</b>	-15.819	1.3485E-07
<b>RAgeID_20_64</b>	-4.512	0.010975388
<b>ID_A00_B99</b>	-12.457	3.89195E-06
<b>ID_C00_D48</b>	3.9337	51.09568238
<b>ID_D50_D89</b>	-12.457	3.89195E-06
<b>ID_E00_E90</b>	-4.4350	0.011855066
<b>ID_F00_F99</b>	-13.1575	1.93095E-06
<b>ID_G00_G99</b>	-12.4642	3.86248E-06
<b>ID_I00_I99</b>	5.198	180.9100597
<b>ID_J00_J99</b>	0.8725	2.392885596
<b>ID_K00_K93</b>	-3.0438	0.047653462
<b>ID_L00_L99</b>	-13.3223	1.63757E-06
<b>ID_M00_M99</b>	-12.743	2.92154E-06
<b>ID_N00_N99</b>	-12.838	2.65916E-06
<b>ID_O00_O99</b>	-8.591	0.000185733
<b>ID_P00_P96</b>	-3.3835	0.033928497
<b>ID_Q00_Q99</b>	-10.1618	3.86177E-05
<b>ID_R00_R99</b>	-12.7384	2.93618E-06

**Región: Biobío**

Tabla 161: *Evaluación - Coeficientes del modelo de RLO de Biobío*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-115.4054	8.1156	-14.22	202.21	<0.0001
<b>y&gt;=3</b>	-117.2229	8.1187	-14.44	208.51	<0.0001
<b>y&gt;=4</b>	-119.2684	8.1237	-14.68	215.50	<0.0001
<b>y&gt;=5</b>	-122.1032	8.1291	-15.02	225.60	<0.0001
<b>TYear</b>	0.0571	0.0040	14.13	199.66	<0.0001
<b>TMonth</b>	0.0361	0.0069	5.20	27.04	<0.0001
<b>DLocationID</b>	-0.8495	0.0494	-17.18	295.15	<0.0001
<b>SexID</b>	-0.4237	0.0484	-8.76	76.74	<0.0001
<b>RAgeID_0_10</b>	-15.2116	48.5348	-0.31	0.10	0.7540
<b>RAgeID_11_19</b>	-7.6374	1.0033	-7.61	57.91	<0.0001
<b>RAgeID_20_64</b>	-2.3592	0.0601	-39.29	1543.70	<0.0001
<b>ID_A00_B99</b>	-14.0634	65.4953	-0.21	0.04	0.8300
<b>ID_C00_D48</b>	2.9409	0.0894	32.88	1081.09	<0.0001
<b>ID_D50_D89</b>	-14.3050	110.0048	-0.13	0.02	0.8965
<b>ID_E00_E90</b>	-3.7430	0.3428	-10.92	119.25	<0.0001
<b>ID_F00_F99</b>	-14.6004	80.6178	-0.18	0.03	0.8563
<b>ID_G00_G99</b>	-4.4908	0.5065	-8.87	78.68	<0.0001
<b>ID_I00_I99</b>	3.1565	0.0911	34.65	1200.62	<0.0001
<b>ID_J00_J99</b>	0.3485	0.0954	3.65	13.32	0.0003
<b>ID_K00_K93</b>	-0.6363	0.1106	-5.75	33.06	<0.0001
<b>ID_L00_L99</b>	-14.7731	131.8465	-0.11	0.01	0.9108
<b>ID_M00_M99</b>	-14.1928	105.9739	-0.13	0.02	0.8935
<b>ID_N00_N99</b>	-4.4738	0.5065	-8.83	77.97	<0.0001
<b>ID_O00_O99</b>	-12.0614	349.0675	-0.03	0.001	0.9724
<b>ID_P00_P96</b>	-5.6563	104.6331	-0.05	0.003	0.9569
<b>ID_Q00_Q99</b>	-11.6319	86.7489	-0.13	0.02	0.8933
<b>ID_R00_R99</b>	-3.6405	0.3633	-10.02	100.40	<0.0001

Tabla 162: *Evaluación - Probabilidades del modelo de RLO de Biobío*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-11.5405	-
<b>y&gt;=3</b>	-11.7223	-
<b>y&gt;=4</b>	-11.9268	-
<b>y&gt;=5</b>	-12.2103	-
<b>TYear</b>	0.0571	1.058761681
<b>TMonth</b>	0.0361	1.036759517
<b>DLocationID</b>	-0.8495	0.427628693
<b>SexID</b>	-0.4237	0.654620239
<b>RAgeID_0_10</b>	-15.212	2.47563E-07
<b>RAgeID_11_19</b>	-7.637	0.00048208
<b>RAgeID_20_64</b>	-2.3592	0.09449579
<b>ID_A00_B99</b>	-14.0634	7.80446E-07
<b>ID_C00_D48</b>	2.9409	18.93287824
<b>ID_D50_D89</b>	-14.3050	6.12939E-07
<b>ID_E00_E90</b>	-3.7430	0.023682948
<b>ID_F00_F99</b>	-14.6004	4.5617E-07
<b>ID_G00_G99</b>	-4.4908	0.011211671
<b>ID_I00_I99</b>	3.1565	23.48824304
<b>ID_J00_J99</b>	0.3485	1.416940543
<b>ID_K00_K93</b>	-0.6363	0.52924702
<b>ID_L00_L99</b>	-14.773	3.83817E-07
<b>ID_M00_M99</b>	-14.193	6.85718E-07
<b>ID_N00_N99</b>	-4.4738	0.011403899
<b>ID_O00_O99</b>	-12.0614	5.77831E-06
<b>ID_P00_P96</b>	-5.6563	0.003495426
<b>ID_Q00_Q99</b>	-11.632	8.8783E-06
<b>ID_R00_R99</b>	-3.6405	0.026239221



**Región: La Araucanía**

Tabla 163: *Evaluación - Coeficientes del modelo de RLO de Araucanía*

<b>Variab</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-73.2836	8.3692	-8.76	76.74	<0.0001
<b>y&gt;=3</b>	-75.0702	8.3709	-8.97	80.46	<0.0001
<b>y&gt;=4</b>	-77.1282	8.3757	-9.21	84.82	<0.0001
<b>y&gt;=5</b>	-79.7584	8.3808	-9.52	90.63	<0.0001
<b>TYear</b>	0.0362	0.0042	8.69	75.52	<0.0001
<b>TMonth</b>	0.0328	0.0072	4.54	20.61	<0.0001
<b>DLocationID</b>	-1.5540	0.0553	-28.08	788.49	<0.0001
<b>SexID</b>	-0.2398	0.0500	-4.79	22.94	<0.0001
<b>RAgeID_0_10</b>	-15.0233	57.7208	-0.26	0.07	0.7947
<b>RAgeID_11_19</b>	-15.6575	74.5691	-0.21	0.04	0.8337
<b>RAgeID_20_64</b>	-2.7688	0.0653	-42.38	1796.06	<0.0001
<b>ID_A00_B99</b>	-14.1018	65.9949	-0.21	0.04	0.8308
<b>ID_C00_D48</b>	2.5169	0.0939	26.82	719.31	<0.0001
<b>ID_D50_D89</b>	-14.3647	124.3953	-0.12	0.01	0.9081
<b>ID_E00_E90</b>	-3.9038	0.3638	-10.73	115.13	<0.0001
<b>ID_F00_F99</b>	-14.6897	81.7830	-0.18	0.03	0.8575
<b>ID_G00_G99</b>	-4.7823	0.5839	-8.19	67.08	<0.0001
<b>ID_I00_I99</b>	2.8251	0.0949	29.77	886.25	<0.0001
<b>ID_J00_J99</b>	0.3398	0.0999	3.40	11.56	0.0007
<b>ID_K00_K93</b>	-2.8074	0.2254	-12.46	155.25	<0.0001
<b>ID_L00_L99</b>	-14.7380	144.1652	-0.10	0.01	0.9186
<b>ID_M00_M99</b>	-14.2619	112.5598	-0.13	0.02	0.8992
<b>ID_N00_N99</b>	-5.8519	1.0038	-5.83	33.99	<0.0001
<b>ID_O00_O99</b>	-11.5841	361.5888	-0.03	0.001	0.9744
<b>ID_P00_P96</b>	0.6001	149.2675	0.00	0.000	0.9668
<b>ID_Q00_Q99</b>	-11.5668	92.8506	-0.12	0.01	0.9009
<b>ID_R00_R99</b>	0.3863	0.1046	3.69	13.62	0.0002

Tabla 164: *Evaluación - Probabilidades del modelo de RLO de Araucanía*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-73.284	-
<b>y&gt;=3</b>	-75.070	-
<b>y&gt;=4</b>	-77.128	-
<b>y&gt;=5</b>	-79.758	-
<b>TYear</b>	0.0362	1.036863198
<b>TMonth</b>	0.0328	1.03334385
<b>DLocationID</b>	-1.554	0.211400678
<b>SexID</b>	-0.2398	0.786785202
<b>RAgeID_0_10</b>	-15.0233	2.98857E-07
<b>RAgeID_11_19</b>	-15.658	1.58502E-07
<b>RAgeID_20_64</b>	-2.7688	0.062737244
<b>ID_A00_B99</b>	-14.1018	7.51045E-07
<b>ID_C00_D48</b>	2.5169	12.39012767
<b>ID_D50_D89</b>	-14.3647	5.77418E-07
<b>ID_E00_E90</b>	-3.9038	0.020165138
<b>ID_F00_F99</b>	-14.6897	4.172E-07
<b>ID_G00_G99</b>	-4.7823	0.00837671
<b>ID_I00_I99</b>	2.8251	16.86263116
<b>ID_J00_J99</b>	0.3398	1.404666629
<b>ID_K00_K93</b>	-2.807	0.060361729
<b>ID_L00_L99</b>	-14.380	5.6865E-07
<b>ID_M00_M99</b>	-14.2619	6.39935E-07
<b>ID_N00_N99</b>	-5.8519	0.002874433
<b>ID_O00_O99</b>	-11.5841	9.31299E-06
<b>ID_P00_P96</b>	0.6001	1.822301021
<b>ID_Q00_Q99</b>	-11.5668	9.47551E-06
<b>ID_R00_R99</b>	0.3863	1.471526062

**Región: Maule**

Tabla 165: *Evaluación - Coeficientes del modelo de RLO de Maule*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-89.0144	9.5559	-9.32	86.86	<0.0001
<b>y&gt;=3</b>	-91.1298	9.5606	-9.53	90.82	<0.0001
<b>y&gt;=4</b>	-93.8679	9.5667	-9.81	96.24	<0.0001
<b>y&gt;=5</b>	-97.1626	9.5841	-10.14	102.82	<0.0001
<b>TYear</b>	0.0440	0.0048	9.25	85.56	<0.0001
<b>TMonth</b>	0.0340	0.0082	4.12	16.97	<0.0001
<b>DLocationID</b>	-1.3133	0.0607	-21.65	468.72	<0.0001
<b>SexID</b>	-0.5796	0.0579	-10.02	100.40	<0.0001
<b>RAgeID_0_10</b>	-6.2370	1.0041	-6.21	38.56	<0.0001
<b>RAgeID_11_19</b>	-15.4443	79.7938	-0.19	0.04	0.8465
<b>RAgeID_20_64</b>	-2.7362	0.0741	-36.90	1361.61	<0.0001
<b>ID_A00_B99</b>	-14.0399	77.7434	-0.18	0.03	0.8567
<b>ID_C00_D48</b>	1.9619	0.0955	20.54	421.89	<0.0001
<b>ID_D50_D89</b>	-14.4375	136.1966	-0.11	0.01	0.9156
<b>ID_E00_E90</b>	-14.3517	69.6988	-0.21	0.04	0.8369
<b>ID_F00_F99</b>	-14.7880	94.7262	-0.16	0.03	0.8759
<b>ID_G00_G99</b>	-14.2315	70.6767	-0.20	0.04	0.8404
<b>ID_I00_I99</b>	2.8146	0.0973	28.92	836.37	<0.0001
<b>ID_J00_J99</b>	-0.0103	0.1065	-0.10	0.01	0.9231
<b>ID_K00_K93</b>	-3.7707	0.3640	-10.36	107.33	<0.0001
<b>ID_L00_L99</b>	-14.8556	170.3893	-0.09	0.01	0.9305
<b>ID_M00_M99</b>	-14.2017	127.7435	-0.11	0.01	0.9115
<b>ID_N00_N99</b>	-14.5315	80.6260	-0.18	0.03	0.8570
<b>ID_O00_O99</b>	-11.3920	461.2778	-0.02	0.000	0.9803
<b>ID_P00_P96</b>	-8.4510	155.6917	-0.05	0.003	0.9567
<b>ID_Q00_Q99</b>	-11.6667	107.9207	-0.11	0.01	0.9139
<b>ID_R00_R99</b>	-4.0832	0.4562	-8.95	80.10	<0.0001

Tabla 166: *Evaluación - Probabilidades del modelo de RLO de Maule*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-89.0144	-
<b>y&gt;=3</b>	-91.1298	-
<b>y&gt;=4</b>	-93.8679	-
<b>y&gt;=5</b>	-97.1626	-
<b>TYear</b>	0.044	1.044982355
<b>TMonth</b>	0.034	1.034584607
<b>DLocationID</b>	-1.313	0.268931118
<b>SexID</b>	-0.5796	0.560122371
<b>RAgeID_0_10</b>	-6.237	1.95571E-03
<b>RAgeID_11_19</b>	-15.4443	1.96167E-07
<b>RAgeID_20_64</b>	-2.7362	0.064816181
<b>ID_A00_B99</b>	-14.0399	7.99004E-07
<b>ID_C00_D48</b>	1.9619	7.112828609
<b>ID_D50_D89</b>	-14.4375	5.36875E-07
<b>ID_E00_E90</b>	-14.3517	5.84973E-07
<b>ID_F00_F99</b>	-14.7880	3.78141E-07
<b>ID_G00_G99</b>	-14.2315	6.59687E-07
<b>ID_I00_I99</b>	2.8146	16.68649985
<b>ID_J00_J99</b>	-0.0103	0.989752863
<b>ID_K00_K93</b>	-3.7707	0.023035932
<b>ID_L00_L99</b>	-14.8556	3.53423E-07
<b>ID_M00_M99</b>	-14.2017	6.79642E-07
<b>ID_N00_N99</b>	-14.5315	4.88708E-07
<b>ID_O00_O99</b>	-11.3920	1.12854E-05
<b>ID_P00_P96</b>	-8.4510	0.000213687
<b>ID_Q00_Q99</b>	-11.667	8.57465E-06
<b>ID_R00_R99</b>	-4.0832	0.016853448

**Región: Libertador General Bernardo O'Higgins**

Tabla 167: *Evaluación - Coeficientes del modelo de RLO de O'Higgins*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-99.2854	9.3625	-10.60	112.36	<0.0001
<b>y&gt;=3</b>	-101.4621	9.3672	-10.83	117.29	<0.0001
<b>y&gt;=4</b>	-103.8713	9.3731	-11.08	122.77	<0.0001
<b>y&gt;=5</b>	-106.3799	9.3808	-11.34	128.60	<0.0001
<b>TYear</b>	0.0492	0.0047	10.55	111.30	<0.0001
<b>TMonth</b>	0.0295	0.0081	3.66	13.40	0.0002
<b>DLocationID</b>	-1.2198	0.0587	-20.80	432.64	<0.0001
<b>SexID</b>	-0.5649	0.0564	-10.01	100.20	<0.0001
<b>RAgeID_0_10</b>	-13.4858	42.9287	-0.31	0.10	0.7534
<b>RAgeID_11_19</b>	-14.3496	55.6059	-0.26	0.07	0.7964
<b>RAgeID_20_64</b>	-2.6305	0.0716	-36.74	1349.83	<0.0001
<b>ID_A00_B99</b>	-13.1306	48.4001	-0.27	0.07	0.7862
<b>ID_C00_D48</b>	1.7072	0.0952	17.94	321.84	<0.0001
<b>ID_D50_D89</b>	-13.4580	84.7601	-0.16	0.03	0.8738
<b>ID_E00_E90</b>	-5.2615	0.7124	-7.39	54.61	<0.0001
<b>ID_F00_F99</b>	-5.7494	1.0040	-5.73	32.83	<0.0001
<b>ID_G00_G99</b>	-13.3416	43.7404	-0.31	0.10	0.7604
<b>ID_I00_I99</b>	2.5763	0.0963	26.76	716.10	<0.0001
<b>ID_J00_J99</b>	0.1982	0.1029	1.93	3.72	0.0540
<b>ID_K00_K93</b>	-2.2441	0.1821	-12.33	152.03	<0.0001
<b>ID_L00_L99</b>	-13.9103	97.4640	-0.14	0.02	0.8865
<b>ID_M00_M99</b>	-13.3059	81.8241	-0.16	0.03	0.8708
<b>ID_N00_N99</b>	-4.6531	0.5842	-7.97	63.52	<0.0001
<b>ID_O00_O99</b>	-10.5762	234.3664	-0.05	0.003	0.9640
<b>ID_P00_P96</b>	-0.0383	105.7518	0.00	0.000	0.9997
<b>ID_Q00_Q99</b>	-10.6484	66.7183	-0.16	0.03	0.8732
<b>ID_R00_R99</b>	-4.4539	0.5080	-8.77	76.91	<0.0001

Tabla 168: *Evaluación - Probabilidades del modelo de RLO de O'Higgins*

<b>Variables</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-99.2854	-
<b>y&gt;=3</b>	-101.4621	-
<b>y&gt;=4</b>	-103.8713	-
<b>y&gt;=5</b>	-106.3799	-
<b>TYear</b>	0.0492	1.050430416
<b>TMonth</b>	0.0295	1.029939435
<b>DLocationID</b>	-1.2198	0.295289219
<b>SexID</b>	-0.5649	0.568416986
<b>RAgeID_0_10</b>	-13.4858	1.39057E-06
<b>RAgeID_11_19</b>	-14.3496	5.86203E-07
<b>RAgeID_20_64</b>	-2.631	0.072042432
<b>ID_A00_B99</b>	-13.131	1.98359E-06
<b>ID_C00_D48</b>	1.70720	5.513502039
<b>ID_D50_D89</b>	-13.4580	1.42977E-06
<b>ID_E00_E90</b>	-5.2615	0.005187518
<b>ID_F00_F99</b>	-5.7494	0.003184691
<b>ID_G00_G99</b>	-13.3416	1.60626E-06
<b>ID_I00_I99</b>	2.5763	13.14839897
<b>ID_J00_J99</b>	0.1982	1.219206211
<b>ID_K00_K93</b>	-2.244	0.106022918
<b>ID_L00_L99</b>	-13.9103	9.09564E-07
<b>ID_M00_M99</b>	-13.3059	1.66464E-06
<b>ID_N00_N99</b>	-4.653	0.009532007
<b>ID_O00_O99</b>	-10.576	2.55161E-05
<b>ID_P00_P96</b>	-0.0383	0.96242417
<b>ID_Q00_Q99</b>	-10.6484	2.37388E-05
<b>ID_R00_R99</b>	-4.4539	0.011633109

## Región: Los Lagos

Tabla 169: *Evaluación - Coeficientes del modelo de RLO de Los Lagos*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-82.7073	9.0174	-9.17	84.09	<0.0001
<b>y&gt;=3</b>	-84.6223	9.0206	-9.38	87.98	<0.0001
<b>y&gt;=4</b>	-87.1053	9.0267	-9.65	93.12	<0.0001
<b>y&gt;=5</b>	-90.3296	9.0388	-9.99	99.80	<0.0001
<b>TYear</b>	0.0410	0.0045	9.13	83.36	<0.0001
<b>TMonth</b>	0.0448	0.0078	5.76	33.18	<0.0001
<b>DLocationID</b>	-1.9713	0.0627	-31.44	988.47	<0.0001
<b>SexID</b>	-0.4030	0.0540	-7.47	55.80	<0.0001
<b>RAgeID_0_10</b>	-14.6026	64.0620	-0.23	0.05	0.8197
<b>RAgeID_11_19</b>	-15.3858	86.0679	-0.18	0.03	0.8581
<b>RAgeID_20_64</b>	-2.2569	0.0663	-34.02	1157.36	<0.0001
<b>ID_A00_B99</b>	-14.2536	72.8525	-0.20	0.04	0.8449
<b>ID_C00_D48</b>	1.9908	0.0948	20.99	440.58	<0.0001
<b>ID_D50_D89</b>	-14.6013	139.9305	-0.10	0.01	0.9169
<b>ID_E00_E90</b>	-4.4587	0.4559	-9.78	95.65	<0.0001
<b>ID_F00_F99</b>	-14.8694	90.7124	-0.16	0.03	0.8698
<b>ID_G00_G99</b>	-5.9477	1.0041	-5.92	35.05	<0.0001
<b>ID_I00_I99</b>	2.3102	0.0950	24.31	590.98	<0.0001
<b>ID_J00_J99</b>	-0.2720	0.1077	-2.52	6.35	0.0116
<b>ID_K00_K93</b>	-2.6029	0.2028	-12.84	164.87	<0.0001
<b>ID_L00_L99</b>	-15.0752	157.4398	-0.10	0.01	0.9237
<b>ID_M00_M99</b>	-14.3926	126.9355	-0.11	0.01	0.9097
<b>ID_N00_N99</b>	-14.6899	80.6394	-0.18	0.03	0.8555
<b>ID_O00_O99</b>	-11.9532	478.0906	-0.03	0.001	0.9801
<b>ID_P00_P96</b>	0.3314	161.0903	0.00	0.000	0.9984
<b>ID_Q00_Q99</b>	-11.9791	102.1740	-0.12	0.01	0.9067
<b>ID_R00_R99</b>	-0.9140	0.1264	-7.23	52.27	<0.0001

Tabla 170: *Evaluación - Probabilidades del modelo de RLO de Los Lagos*

<b>Variabes</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-82.7073	-
<b>y&gt;=3</b>	-84.6223	-
<b>y&gt;=4</b>	-87.1053	-
<b>y&gt;=5</b>	-90.3296	-
<b>TYear</b>	0.0410	1.041852106
<b>TMonth</b>	0.0448	1.045818675
<b>DLocationID</b>	-1.9713	0.13927568
<b>SexID</b>	-0.4030	0.668312099
<b>RAgeID_0_10</b>	-14.6026	4.55168E-07
<b>RAgeID_11_19</b>	-15.3858	2.07985E-07
<b>RAgeID_20_64</b>	-2.257	0.104674473
<b>ID_A00_B99</b>	-14.2536	6.45268E-07
<b>ID_C00_D48</b>	1.9908	7.321388531
<b>ID_D50_D89</b>	-14.6013	4.5576E-07
<b>ID_E00_E90</b>	-4.4587	0.011577404
<b>ID_F00_F99</b>	-14.8694	3.48579E-07
<b>ID_G00_G99</b>	-5.948	0.002611841
<b>ID_I00_I99</b>	2.3102	10.07643974
<b>ID_J00_J99</b>	-0.272	0.761854261
<b>ID_K00_K93</b>	-2.6029	0.074058497
<b>ID_L00_L99</b>	-15.0752	2.83742E-07
<b>ID_M00_M99</b>	-14.3926	5.6153E-07
<b>ID_N00_N99</b>	-14.6899	4.17117E-07
<b>ID_O00_O99</b>	-11.9532	6.4386E-06
<b>ID_P00_P96</b>	0.3314	1.392916848
<b>ID_Q00_Q99</b>	-11.9791	6.27398E-06
<b>ID_R00_R99</b>	-0,914	0,400917343



## Región: Coquimbo

Tabla 171: *Evaluación - Coeficientes del modelo de RLO de Coquimbo*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-144.9490	9.9777	-14.53	211.12	<0.0001
<b>y&gt;=3</b>	-147.1540	9.9836	-14.74	217.27	<0.0001
<b>y&gt;=4</b>	-149.0705	9.9919	-14.92	222.61	<0.0001
<b>y&gt;=5</b>	-151.0372	9.9970	-15.11	228.31	<0.0001
<b>TYear</b>	0.0717	0.0050	14.44	208.51	<0.0001
<b>TMonth</b>	0.0360	0.0084	4.27	18.23	<0.0001
<b>DLocationID</b>	-2.1366	0.0707	-30.20	912.04	<0.0001
<b>SexID</b>	-0.2974	0.0586	-5.08	25.81	<0.0001
<b>RAgeID_0_10</b>	-14.6509	61.2285	-0.24	0.06	0.8109
<b>RAgeID_11_19</b>	-15.5068	88.6504	-0.17	0.03	0.8611
<b>RAgeID_20_64</b>	-3.2163	0.0829	-38.80	1505.44	<0.0001
<b>ID_A00_B99</b>	-13.7314	70.7076	-0.19	0.04	0.8460
<b>ID_C00_D48</b>	3.2188	0.1214	26.52	703.31	<0.0001
<b>ID_D50_D89</b>	-14.1500	154.6021	-0.09	0.01	0.9271
<b>ID_E00_E90</b>	-2.8162	0.2896	-9.72	94.48	<0.0001
<b>ID_F00_F99</b>	-14.5669	109.0604	-0.13	0.02	0.8937
<b>ID_G00_G99</b>	-4.2288	0.5885	-7.19	51.70	<0.0001
<b>ID_I00_I99</b>	3.4916	0.1224	28.53	813.96	<0.0001
<b>ID_J00_J99</b>	0.2554	0.1297	1.97	3.88	0.0490
<b>ID_K00_K93</b>	-2.6900	0.2813	-9.56	91.39	<0.0001
<b>ID_L00_L99</b>	-14.6720	193.3180	-0.08	0.01	0.9395
<b>ID_M00_M99</b>	-14.0517	132.0127	-0.11	0.01	0.9152
<b>ID_N00_N99</b>	-14.1626	82.2854	-0.17	0.03	0.8633
<b>ID_O00_O99</b>	-10.2915	400.8525	-0.03	0.001	0.9795
<b>ID_P00_P96</b>	-6.3857	124.8502	-0.05	0.003	0.9592
<b>ID_Q00_Q99</b>	-10.6687	104.6765	-0.10	0.01	0.9188
<b>ID_R00_R99</b>	-1.3403	0.1758	-7.62	58.06	<0.0001

Tabla 172: *Evaluación - Probabilidades del modelo de RLO de Coquimbo*

<b>Variables</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-144.9490	-
<b>y&gt;=3</b>	-147.1540	-
<b>y&gt;=4</b>	-149.0705	-
<b>y&gt;=5</b>	-151.0372	-
<b>TYear</b>	0.0717	1.074332996
<b>TMonth</b>	0.036	1.036655846
<b>DLocationID</b>	-2.137	0.11805555
<b>SexID</b>	-0.2974	0.742746854
<b>RAgeID_0_10</b>	-14.651	4.33706E-07
<b>RAgeID_11_19</b>	-15.5068	1.84282E-07
<b>RAgeID_20_64</b>	-3.2163	0.040103166
<b>ID_A00_B99</b>	-13.7314	1.08775E-06
<b>ID_C00_D48</b>	3.2188	24.99810445
<b>ID_D50_D89</b>	-14.1500	7.15703E-07
<b>ID_E00_E90</b>	-2.8162	0.059832876
<b>ID_F00_F99</b>	-14.567	4.71711E-07
<b>ID_G00_G99</b>	-4.2288	0.014569864
<b>ID_I00_I99</b>	3.4916	32.83844721
<b>ID_J00_J99</b>	0.2554	1.290977909
<b>ID_K00_K93</b>	-2.690	0.067880939
<b>ID_L00_L99</b>	-14.672	4.2465E-07
<b>ID_M00_M99</b>	-14.052	7.89631E-07
<b>ID_N00_N99</b>	-14.163	7.06742E-07
<b>ID_O00_O99</b>	-10.2915	3.39202E-05
<b>ID_P00_P96</b>	-6.3857	0.001685488
<b>ID_Q00_Q99</b>	-10.6687	2.32618E-05
<b>ID_R00_R99</b>	-13.4030	1.51061E-06

**Región: Ñuble**

Tabla 173: *Evaluación - Coeficientes del modelo de RLO de Ñuble*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-69.6647	9.1431	-7.62	58.06	<0.0001
<b>y&gt;=3</b>	-71.5687	9.1459	-7.83	61.31	<0.0001
<b>y&gt;=4</b>	-73.7295	9.1491	-8.06	64.96	<0.0001
<b>y&gt;=5</b>	-76.2137	9.1539	-8.33	69.39	<0.0001
<b>TYear</b>	0.0344	0.0046	7.56	57.15	<0.0001
<b>TMonth</b>	0.0210	0.0079	2.67	7.13	0.0077
<b>DLocationID</b>	-1.0691	0.0572	-18.70	349.69	<0.0001
<b>SexID</b>	-0.4888	0.0553	-8.83	77.97	<0.0001
<b>RAgeID_0_10</b>	-13.3397	50.2713	-0.27	0.07	0.7907
<b>RAgeID_11_19</b>	-14.0326	66.7534	-0.21	0.04	0.8335
<b>RAgeID_20_64</b>	-2.1999	0.0671	-32.76	1073.22	<0.0001
<b>ID_A00_B99</b>	-13.0739	51.6123	-0.25	0.06	0.8000
<b>ID_C00_D48</b>	1.5989	0.0943	16.96	287.64	<0.0001
<b>ID_D50_D89</b>	-13.4263	93.9755	-0.14	0.02	0.8864
<b>ID_E00_E90</b>	-3.3005	0.3016	-10.94	119.68	<0.0001
<b>ID_F00_F99</b>	-4.3864	0.5845	-7.50	56.25	<0.0001
<b>ID_G00_G99</b>	-5.4508	1.0042	-5.43	29.48	<0.0001
<b>ID_I00_I99</b>	2.5133	0.0952	26.40	696.96	<0.0001
<b>ID_J00_J99</b>	-0.4207	0.1126	-3.74	13.99	0.0002
<b>ID_K00_K93</b>	-1.1980	0.1354	-8.85	78.32	<0.0001
<b>ID_L00_L99</b>	-13.7946	111.5595	-0.12	0.01	0.9016
<b>ID_M00_M99</b>	-13.2001	99.1504	-0.13	0.02	0.8941
<b>ID_N00_N99</b>	-5.4714	1.0040	-5.45	29.70	<0.0001
<b>ID_O00_O99</b>	-10.8562	341.8506	-0.03	0.001	0.9747
<b>ID_P00_P96</b>	-0.1237	117.5302	0.00	0.000	0.9992
<b>ID_Q00_Q99</b>	-10.4934	73.8935	-0.14	0.02	0.8871
<b>ID_R00_R99</b>	-3.7370	0.4185	-8.93	79.74	<0.0001

Tabla 174: *Evaluación - Probabilidades del modelo de RLO de Nuble*

<b>Variabes</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-69.6647	-
<b>y&gt;=3</b>	-71.5687	-
<b>y&gt;=4</b>	-73.7295	-
<b>y&gt;=5</b>	-76.2137	-
<b>TYear</b>	0.0344	1.092097326
<b>TMonth</b>	0.021	1.025725329
<b>DLocationID</b>	-1.069	0.098136099
<b>SexID</b>	-0.4888	1.153153081
<b>RAgeID_0_10</b>	-13.340	7.47737E-08
<b>RAgeID_11_19</b>	-14.033	3.29393E-08
<b>RAgeID_20_64</b>	-2.200	0.017653875
<b>ID_A00_B99</b>	-13.074	4.90618E-07
<b>ID_C00_D48</b>	1.599	21.86090759
<b>ID_D50_D89</b>	-13.426	3.29003E-07
<b>ID_E00_E90</b>	-3.301	0.00894107
<b>ID_F00_F99</b>	-4.386	0.004508909
<b>ID_G00_G99</b>	-5.451	6.37507E-07
<b>ID_I00_I99</b>	2.513	64.71545211
<b>ID_J00_J99</b>	-0.4207	1.109711421
<b>ID_K00_K93</b>	-1.198	4.55168E-07
<b>ID_L00_L99</b>	-13.795	2.01677E-07
<b>ID_M00_M99</b>	-13.200	3.4851E-07
<b>ID_N00_N99</b>	-5.471	3.73107E-07
<b>ID_O00_O99</b>	-10.856	1.30712E-05
<b>ID_P00_P96</b>	-0.1237	6.41397E-35
<b>ID_Q00_Q99</b>	-10.493	2.54291E-06
<b>ID_R00_R99</b>	-3.737	5.04549E-07

**Región: Antofagasta**

Tabla 175: *Evaluación - Coeficientes del modelo de RLO de Antofagasta*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-97.2407	9.3727	-10.37	107.54	<0.0001
<b>y&gt;=3</b>	-99.5974	9.3775	-10.62	112.78	<0.0001
<b>y&gt;=4</b>	-101.6811	9.3811	-10.84	117.51	<0.0001
<b>y&gt;=5</b>	-104.2783	9.3893	-11.11	123.43	<0.0001
<b>TYear</b>	0.0481	0.0047	10.31	106.30	<0.0001
<b>TMonth</b>	0.0159	0.0081	1.97	3.88	<0.0001
<b>DLocationID</b>	-0.9760	0.0576	-16.93	286.62	0.0489
<b>SexID</b>	-0.7321	0.0571	-12.83	164.61	<0.0001
<b>RAgeID_0_10</b>	-6.0189	1.0870	-5.54	30.69	<0.0001
<b>RAgeID_11_19</b>	-13.1531	36.5606	-0.36	0.13	<0.0001
<b>RAgeID_20_64</b>	-1.5952	0.0615	-25.95	673.40	0.7190
<b>ID_A00_B99</b>	-4.8910	1.0048	-4.87	23.72	<0.0001
<b>ID_C00_D48</b>	2.3604	0.0983	24.02	576.96	<0.0001
<b>ID_D50_D89</b>	-12.2422	63.5813	-0.19	0.04	<0.0001
<b>ID_E00_E90</b>	-3.1743	0.3036	-10.46	109.41	0.8473
<b>ID_F00_F99</b>	-2.6947	0.2763	-9.75	95.06	<0.0001
<b>ID_G00_G99</b>	-3.6516	0.4581	-7.97	63.52	<0.0001
<b>ID_I00_I99</b>	2.0547	0.0981	20.95	438.90	<0.0001
<b>ID_J00_J99</b>	-0.8505	0.1285	-6.62	43.82	<0.0001
<b>ID_K00_K93</b>	-1.7373	0.1714	-10.13	102.62	<0.0001
<b>ID_L00_L99</b>	-12.3929	77.7519	-0.16	0.03	0.8734
<b>ID_M00_M99</b>	-12.0236	58.6970	-0.20	0.04	0.8377
<b>ID_N00_N99</b>	-4.7025	0.7136	-6.59	43.43	<0.0001
<b>ID_O00_O99</b>	-10.2942	152.5669	-0.07	0.005	0.9462
<b>ID_P00_P96</b>	0.6803	1.1548	0.59	0.348	0.5558
<b>ID_Q00_Q99</b>	-9.4915	41.6516	-0.23	0.05	0.8197
<b>ID_R00_R99</b>	-4.5983	0.7140	-6.44	41.47	<0.0001

Tabla 176: *Evaluación - Probabilidades del modelo de RLO de Antofagasta*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-97.2407	-
<b>y&gt;=3</b>	-99.5974	-
<b>y&gt;=4</b>	-101.6811	-
<b>y&gt;=5</b>	-104.2783	-
<b>TYear</b>	0.0481	1.049275578
<b>TMonth</b>	0.0159	1.016027078
<b>DLocationID</b>	-0.976	0.37681535
<b>SexID</b>	-0.7321	0.480898043
<b>RAgeID_0_10</b>	-6.019	2.43234E-03
<b>RAgeID_11_19</b>	-13.153	1.93946E-06
<b>RAgeID_20_64</b>	-1.595	0.202867951
<b>ID_A00_B99</b>	-4.891	0.007513905
<b>ID_C00_D48</b>	2.360	10.59518868
<b>ID_D50_D89</b>	-12.242	4.82259E-06
<b>ID_E00_E90</b>	-3.174	0.04182337
<b>ID_F00_F99</b>	-2.695	0.067562648
<b>ID_G00_G99</b>	-3.652	0.025949576
<b>ID_I00_I99</b>	2.055	7.804496173
<b>ID_J00_J99</b>	-0.851	0.427201278
<b>ID_K00_K93</b>	-1.737	0.175994946
<b>ID_L00_L99</b>	-12.393	4.14793E-06
<b>ID_M00_M99</b>	-12.024	6.00091E-06
<b>ID_N00_N99</b>	-4.703	0.009072567
<b>ID_O00_O99</b>	-10.294	3.38287E-05
<b>ID_P00_P96</b>	0.6803	1.974469984
<b>ID_Q00_Q99</b>	-9.492	7.54908E-05
<b>ID_R00_R99</b>	-4.598	0.010068938

## Región: Los Ríos

Tabla 177: *Evaluación - Coeficientes del modelo de RLO de Los Ríos*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-63.7180	8.8123	-7.23	52.27	<0.0001
<b>y&gt;=3</b>	-65.5672	8.8146	-7.44	55.35	<0.0001
<b>y&gt;=4</b>	-67.4980	8.8179	-7.65	58.52	<0.0001
<b>y&gt;=5</b>	-69.6763	8.8211	-7.90	62.41	<0.0001
<b>TYear</b>	0.0316	0.0044	7.19	51.70	<0.0001
<b>TMonth</b>	0.0312	0.0077	4.06	16.48	<0.0001
<b>DLocationID</b>	-1.0534	0.0555	-18.99	360.62	<0.0001
<b>SexID</b>	-0.5232	0.0539	-9.72	94.48	<0.0001
<b>RAgeID_0_10</b>	-12.5314	30.4802	-0.41	0.17	0.6810
<b>RAgeID_11_19</b>	-4.7703	0.5857	-8.14	66.26	<0.0001
<b>RAgeID_20_64</b>	-2.3490	0.0665	-35.30	1246.09	<0.0001
<b>ID_A00_B99</b>	-5.3600	1.0045	-5.34	28.52	<0.0001
<b>ID_C00_D48</b>	1.7118	0.0953	17.96	322.56	<0.0001
<b>ID_D50_D89</b>	-12.4584	69.5670	-0.18	0.03	0.8579
<b>ID_E00_E90</b>	-3.0759	0.2524	-12.19	148.60	<0.0001
<b>ID_F00_F99</b>	-5.0212	0.7134	-7.04	49.56	<0.0001
<b>ID_G00_G99</b>	-3.6885	0.3657	-10.09	101.81	<0.0001
<b>ID_I00_I99</b>	2.3914	0.0965	24.78	614.05	<0.0001
<b>ID_J00_J99</b>	-0.0657	0.1049	-0.63	0.40	0.5308
<b>ID_K00_K93</b>	-2.4601	0.1981	-12.42	154.26	<0.0001
<b>ID_L00_L99</b>	-12.8413	65.9557	-0.19	0.04	0.8456
<b>ID_M00_M99</b>	-12.4004	56.1631	-0.22	0.05	0.8253
<b>ID_N00_N99</b>	-3.9892	0.4185	-9.53	90.82	<0.0001
<b>ID_O00_O99</b>	-9.7193	194.2381	-0.05	0.003	0.9601
<b>ID_P00_P96</b>	-0.0976	70.4202	0.00	0.000	0.9989
<b>ID_Q00_Q99</b>	-9.2249	44.9778	-0.21	0.04	0.8375
<b>ID_R00_R99</b>	-1.929	0.1663	-11.60	134.56	<0.0001

Tabla 178: *Evaluación - Probabilidades del modelo de RLO de Los Ríos*

<b>Variabes</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-637.180	-
<b>y&gt;=3</b>	-655.672	-
<b>y&gt;=4</b>	-674.980	-
<b>y&gt;=5</b>	-696.763	-
<b>TYear</b>	0.0316	1.032104581
<b>TMonth</b>	0.0312	1.031691822
<b>DLocationID</b>	-1.053	0.348749981
<b>SexID</b>	-0.5232	0.592621123
<b>RAgeID_0_10</b>	-12.531	3.61145E-06
<b>RAgeID_11_19</b>	-4.770	0.008477836
<b>RAgeID_20_64</b>	-2.349	0.095464579
<b>ID_A00_B99</b>	-5.360	0.004700906
<b>ID_C00_D48</b>	1.712	5.538922571
<b>ID_D50_D89</b>	-12.458	3.88495E-06
<b>ID_E00_E90</b>	-3.076	0.046148076
<b>ID_F00_F99</b>	-5.021	0.006596606
<b>ID_G00_G99</b>	-3.689	0.025009488
<b>ID_I00_I99</b>	2.391	10.92878353
<b>ID_J00_J99</b>	-0.0657	0.936411746
<b>ID_K00_K93</b>	-2.460	0.085426408
<b>ID_L00_L99</b>	-12.841	2.64908E-06
<b>ID_M00_M99</b>	-12.400	4.11694E-06
<b>ID_N00_N99</b>	-3.989	0.01851452
<b>ID_O00_O99</b>	-9.719	6.01121E-05
<b>ID_P00_P96</b>	-0.0976	0.907011636
<b>ID_Q00_Q99</b>	-9.225	9.85546E-05
<b>ID_R00_R99</b>	-1.929	0.145293419



**Región: Atacama**

Tabla 179: *Evaluación - Coeficientes del modelo de RLO de Atacama*

<b>Variabes</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-98.3291	10.9551	-8.98	80.64	<0.0001
<b>y&gt;=3</b>	-100.6401	10.9608	-9.18	84.27	<0.0001
<b>y&gt;=4</b>	-102.8853	10.9643	-9.38	87.98	<0.0001
<b>y&gt;=5</b>	-106.3686	11.0089	-9.66	93.32	<0.0001
<b>TYear</b>	0.0486	0.0055	8.90	79.21	<0.0001
<b>TMonth</b>	0.0329	0.0095	3.47	12.04	0.0005
<b>DLocationID</b>	-0.7209	0.0665	-10.84	117.51	<0.0001
<b>SexID</b>	-0.5351	0.0667	-8.02	64.32	<0.0001
<b>RAgeID_0_10</b>	-4.0082	0.7244	-5.53	30.58	<0.0001
<b>RAgeID_11_19</b>	-4.7631	1.0061	-4.73	22.37	<0.0001
<b>RAgeID_20_64</b>	-2.0867	0.0819	-25.49	649.7401	<0.0001
<b>ID_A00_B99</b>	-4.6061	1.0075	-4.57	20.88	<0.0001
<b>ID_C00_D48</b>	1.0634	0.1199	8.87	78.68	<0.0001
<b>ID_D50_D89</b>	-12.3374	93.2121	-0.13	0.02	0.8947
<b>ID_E00_E90</b>	-3.2321	0.3722	-8.68	75.34	<0.0001
<b>ID_F00_F99</b>	-12.7504	58.3275	-0.22	0.05	0.8270
<b>ID_G00_G99</b>	-3.8360	0.5902	-6.50	42.25	<0.0001
<b>ID_I00_I99</b>	1.4847	0.1193	12.44	154.75	<0.0001
<b>ID_J00_J99</b>	-0.3378	0.1365	-2.48	6.15	0.0133
<b>ID_K00_K93</b>	-2.2321	0.2465	-9.06	82.08	<0.0001
<b>ID_L00_L99</b>	-12.7939	121.7150	-0.11	0.01	0.9163
<b>ID_M00_M99</b>	-12.0675	79.7754	-0.15	0.02	0.8798
<b>ID_N00_N99</b>	-12.5460	43.0062	-0.29	0.08	0.7705
<b>ID_O00_O99</b>	-10.2307	295.2491	-0.03	0.00	0.9724
<b>ID_P00_P96</b>	1.2123	0.8850	1.37	1.877	0.1707
<b>ID_Q00_Q99</b>	-1.0774	0.7504	-1.44	2.074	0.1510
<b>ID_R00_R99</b>	-2.9145	0.3545	-8.22	67.57	<0.0001

Tabla 180: *Evaluación - Probabilidades del modelo de RLO de Atacama*

<b>Variab</b> les	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-98.3291	-
<b>y&gt;=3</b>	-100.6401	-
<b>y&gt;=4</b>	-102.8853	-
<b>y&gt;=5</b>	-106.3686	-
<b>TYear</b>	0.0486	1.049800347
<b>TMonth</b>	0.0329	1.033447189
<b>DLocationID</b>	-0.7209	0.486314376
<b>SexID</b>	-0.5351	0.585610726
<b>RAgeID_0_10</b>	-4.008	1.81661E-02
<b>RAgeID_11_19</b>	-4.763	0.008539097
<b>RAgeID_20_64</b>	-2.0867	0.124095978
<b>ID_A00_B99</b>	-4.606	0.009990706
<b>ID_C00_D48</b>	1.063	2.896201353
<b>ID_D50_D89</b>	-12.337	4.38465E-06
<b>ID_E00_E90</b>	-3.232	0.039474515
<b>ID_F00_F99</b>	-12.750	2.90116E-06
<b>ID_G00_G99</b>	-3.836	0.021579748
<b>ID_I00_I99</b>	1.485	4.413641122
<b>ID_J00_J99</b>	-0.3378	0.713337941
<b>ID_K00_K93</b>	-2.232	0.107302857
<b>ID_L00_L99</b>	-12.794	2.77766E-06
<b>ID_M00_M99</b>	-12.068	5.74317E-06
<b>ID_N00_N99</b>	-12.546	3.55911E-06
<b>ID_O00_O99</b>	-10.231	3.60465E-05
<b>ID_P00_P96</b>	1.212	3.361206544
<b>ID_Q00_Q99</b>	-1.077	0.340479623
<b>ID_R00_R99</b>	-2.915	0.05423114

**Región: Tarapacá**

Tabla 181: *Evaluación - Coeficientes del modelo de RLO de Tarapacá*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-134.5062	12.3845	-10.86	117.94	<0.0001
<b>y&gt;=3</b>	-136.8749	12.3917	-11.05	122.10	<0.0001
<b>y&gt;=4</b>	-139.3545	12.3968	-11.24	126.34	<0.0001
<b>y&gt;=5</b>	-141.4448	12.4148	-11.39	129.73	<0.0001
<b>TYear</b>	0.0665	0.0062	10.79	116.42	<0.0001
<b>TMonth</b>	0.0384	0.0105	3.66	13.40	0.0002
<b>DLocationID</b>	-1.2185	0.0771	-15.81	249.96	<0.0001
<b>SexID</b>	-0.4764	0.0737	-6.46	41.73	<0.0001
<b>RAgeID_0_10</b>	-4.1818	1.0265	-4.07	16.56	<0.0001
<b>RAgeID_11_19</b>	-4.4885	1.0065	-4.46	19.89	<0.0001
<b>RAgeID_20_64</b>	-1.4423	0.0812	-17.77	315.77	<0.0001
<b>ID_A00_B99</b>	-3.7495	0.7174	-5.23	27.35	<0.0001
<b>ID_C00_D48</b>	0.8564	0.1181	7.25	52.56	<0.0001
<b>ID_D50_D89</b>	-11.6332	80.7630	-0.14	0.02	0.8855
<b>ID_E00_E90</b>	-5.1882	1.0068	-5.15	26.52	<0.0001
<b>ID_F00_F99</b>	-3.3448	0.4636	-7.22	52.13	<0.0001
<b>ID_G00_G99</b>	-4.2664	0.7176	-5.95	35.40	<0.0001
<b>ID_I00_I99</b>	1.0559	0.1177	8.97	80.46	<0.0001
<b>ID_J00_J99</b>	-1.9254	0.2080	-9.26	85.75	<0.0001
<b>ID_K00_K93</b>	-1.8582	0.2174	-8.55	73.10	<0.0001
<b>ID_L00_L99</b>	-11.6682	78.6453	-0.15	0.02	0.8821
<b>ID_M00_M99</b>	-11.4982	53.4476	-0.22	0.05	0.8297
<b>ID_N00_N99</b>	-4.1205	0.7175	-5.74	32.95	<0.0001
<b>ID_O00_O99</b>	-10.0822	167.8085	-0.06	0.004	0.9521
<b>ID_P00_P96</b>	-7.2806	50.4525	-0.14	0.020	0.8853
<b>ID_Q00_Q99</b>	-1.0316	1.0429	-0.99	0.98	0.3226
<b>ID_R00_R99</b>	-4.8835	1.0075	-4.85	23.52	<0.0001

Tabla 182: *Evaluación - Probabilidades del modelo de RLO de Tarapacá*

<b>Variables</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-1.345.062	-
<b>y&gt;=3</b>	-1.368.749	-
<b>y&gt;=4</b>	-1.393.545	-
<b>y&gt;=5</b>	-1.414.448	-
<b>TYear</b>	0.0665	1.068760964
<b>TMonth</b>	0.0384	1.039146808
<b>DLocationID</b>	-1.219	0.295673344
<b>SexID</b>	-0.4764	0.621015027
<b>RAgeID_0_10</b>	-4.182	1.52710E-02
<b>RAgeID_11_19</b>	-4.489	0.011237487
<b>RAgeID_20_64</b>	-1.442	0.236383451
<b>ID_A00_B99</b>	-3.750	0.023529508
<b>ID_C00_D48</b>	0.8564	2.35466861
<b>ID_D50_D89</b>	-11.633	8.86677E-06
<b>ID_E00_E90</b>	-5.188	0.005582045
<b>ID_F00_F99</b>	-3.345	0.035267268
<b>ID_G00_G99</b>	-4.266	0.014032208
<b>ID_I00_I99</b>	1.056	2.874561095
<b>ID_J00_J99</b>	-1.925	0.145817418
<b>ID_K00_K93</b>	-1.858	0.155953093
<b>ID_L00_L99</b>	-11.668	8.5618E-06
<b>ID_M00_M99</b>	-11.498	1.01483E-05
<b>ID_N00_N99</b>	-4.121	0.016236394
<b>ID_O00_O99</b>	-10.082	4.18173E-05
<b>ID_P00_P96</b>	-7.281	0.000688772
<b>ID_Q00_Q99</b>	-1.032	0.356436206
<b>ID_R00_R99</b>	-4.884	0.007570471

**Región: Arica y Parinacota**

Tabla 183: *Evaluación - Coeficientes del modelo de RLO de Arica y Parinacota*

<b>Variab</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-97.6005	11.0306	-8.85	78.32	<0.0001
<b>y&gt;=3</b>	-99.6466	11.0349	-9.03	81.54	<0.0001
<b>y&gt;=4</b>	-101.2607	11.0380	-9.17	84.09	<0.0001
<b>y&gt;=5</b>	-102.9714	11.0418	-9.33	87.05	<0.0001
<b>TYear</b>	0.0482	0.0055	8.78	77.09	<0.0001
<b>TMonth</b>	0.0238	0.0094	2.52	6.35	0.0117
<b>DLocationID</b>	-1.3153	0.0708	-18.57	344.84	<0.0001
<b>SexID</b>	-0.2797	0.0658	-4.25	18.06	<0.0001
<b>RAgeID_0_10</b>	-4.1660	1.0116	-4.12	16.97	<0.0001
<b>RAgeID_11_19</b>	-4.4352	1.0100	-4.39	19.27	<0.0001
<b>RAgeID_20_64</b>	-1.7061	0.0786	-21.71	471.32	<0.0001
<b>ID_A00_B99</b>	-3.2046	0.5144	-6.23	38.81	<0.0001
<b>ID_C00_D48</b>	1.5116	0.1172	12.89	166.15	<0.0001
<b>ID_D50_D89</b>	-11.1232	63.9490	-0.17	0.03	0.8619
<b>ID_E00_E90</b>	-1.7168	0.2024	-8.48	71.91	<0.0001
<b>ID_F00_F99</b>	-2.8614	0.4274	-6.70	44.89	<0.0001
<b>ID_G00_G99</b>	-3.3215	0.4636	-7.16	51.27	<0.0001
<b>ID_I00_I99</b>	1.2088	0.1185	10.20	104.04	<0.0001
<b>ID_J00_J99</b>	-0.8634	0.1529	-5.65	31.92	<0.0001
<b>ID_K00_K93</b>	-1.2910	0.1810	-7.13	50.84	<0.0001
<b>ID_L00_L99</b>	-11.6804	82.7742	-0.14	0.02	0.8878
<b>ID_M00_M99</b>	-11.2634	45.9355	-0.25	0.06	0.8063
<b>ID_N00_N99</b>	-2.7150	0.3380	-8.03	64.48	<0.0001
<b>ID_O00_O99</b>	-9.2356	138.6561	-0.07	0.005	0.9469
<b>ID_P00_P96</b>	-6.8839	51.4774	-0.13	0.017	0.8936
<b>ID_Q00_Q99</b>	-1.7131	1.0402	-1.65	2.72	0.0996
<b>ID_R00_R99</b>	-2.7512	0.3551	-7.75	60.06	<0.0001

Tabla 184: Evaluación - Probabilidades del modelo de RLO de Arica y Parinacota

<b>Variables</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-976.005	-
<b>y&gt;=3</b>	-996.466	-
<b>y&gt;=4</b>	-1.012.607	-
<b>y&gt;=5</b>	-1.029.714	-
<b>TYear</b>	0.0482	-976.005
<b>TMonth</b>	0.0238	-996.466
<b>DLocationID</b>	-1.315	-1.012.607
<b>SexID</b>	-0.2797	-1.029.714
<b>RAgeID_0_10</b>	-4.166	0.0482
<b>RAgeID_11_19</b>	-4.435	0.0238
<b>RAgeID_20_64</b>	-1.706	-1.315
<b>ID_A00_B99</b>	-3.205	-0.2797
<b>ID_C00_D48</b>	1.512	-4.166
<b>ID_D50_D89</b>	-11.123	-4.435
<b>ID_E00_E90</b>	-1.717	-1.706
<b>ID_F00_F99</b>	-2.861	-3.205
<b>ID_G00_G99</b>	-3.322	1.512
<b>ID_I00_I99</b>	1.209	-11.123
<b>ID_J00_J99</b>	-0.8634	-1.717
<b>ID_K00_K93</b>	-1.291	-2.861
<b>ID_L00_L99</b>	-11.680	-3.322
<b>ID_M00_M99</b>	-11.263	1.209
<b>ID_N00_N99</b>	-2.715	-0.8634
<b>ID_O00_O99</b>	-9.236	-1.291
<b>ID_P00_P96</b>	-6.884	-11.680
<b>ID_Q00_Q99</b>	-1.713	-11.263
<b>ID_R00_R99</b>	-976.005	-2.715

## Región: Magallanes y de la Antártica Chilena

Tabla 185: *Evaluación - Coeficientes del modelo de RLO de Magallanes*

<b>Variabes</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-44.2011	10.8101	-4.09	16.73	<0.0001
<b>y&gt;=3</b>	-46.3751	10.8129	-4.29	18.40	<0.0001
<b>y&gt;=4</b>	-48.5992	10.8149	-4.49	20.16	<0.0001
<b>y&gt;=5</b>	-50.8950	10.8254	-4.70	22.09	<0.0001
<b>TYear</b>	0.0216	0.0054	4.02	16.16	<0.0001
<b>TMonth</b>	0.0168	0.0095	1.77	3.13	0.0773
<b>DLocationID</b>	-0.3366	0.0653	-5.15	26.52	<0.0001
<b>SexID</b>	-0.6055	0.0672	-9.02	81.36	<0.0001
<b>RAgeID_0_10</b>	-11.1823	42.6633	-0.26	0.07	0.7932
<b>RAgeID_11_19</b>	-12.9035	64.7113	-0.20	0.04	0.8419
<b>RAgeID_20_64</b>	-1.8094	0.0775	-23.34	544.76	<0.0001
<b>ID_A00_B99</b>	-12.2600	50.6330	-0.24	0.06	0.8087
<b>ID_C00_D48</b>	0.9495	0.1200	7.91	62.57	<0.0001
<b>ID_D50_D89</b>	-12.3157	102.0596	-0.12	0.01	0.9040
<b>ID_E00_E90</b>	-2.7569	0.3375	-8.17	66.75	<0.0001
<b>ID_F00_F99</b>	-3.6375	0.5901	-6.16	37.95	<0.0001
<b>ID_G00_G99</b>	12.2905	45.2287	-0.27	0.07	0.7858
<b>ID_I00_I99</b>	1.6390	0.1189	13.79	190.16	<0.0001
<b>ID_J00_J99</b>	-0.9170	0.1575	-5.82	33.87	<0.0001
<b>ID_K00_K93</b>	-1.3383	0.1847	-7.25	52.56	<0.0001
<b>ID_L00_L99</b>	-12.5386	112.5177	-0.11	0.01	0.9113
<b>ID_M00_M99</b>	-12.1001	88.0928	-0.14	0.02	0.8907
<b>ID_N00_N99</b>	-2.8872	0.3970	-7.27	52.85	<0.0001
<b>ID_O00_O99</b>	-10.6108	374.8330	-0.03	0.001	0.9774
<b>ID_P00_P96</b>	-5.5047	70.3872	-0.08	0.006	0.9377
<b>ID_Q00_Q99</b>	-8.9693	66.7229	-0.13	0.02	0.8931
<b>ID_R00_R99</b>	-12.2132	58.8241	-0.21	0.04	0.8355

Tabla 186: *Evaluación - Probabilidades del modelo de RLO de Magallanes*

<b>Variables</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-442.011	-
<b>y&gt;=3</b>	-463.751	-
<b>y&gt;=4</b>	-485.992	-
<b>y&gt;=5</b>	-508.950	-
<b>TYear</b>	0.0216	1.021834969
<b>TMonth</b>	0.0168	1.016941914
<b>DLocationID</b>	-0.3366	0.714194461
<b>SexID</b>	-0.6055	0.545801458
<b>RAgeID_0_10</b>	-11.182	1.39184E-05
<b>RAgeID_11_19</b>	-12.904	2.48932E-06
<b>RAgeID_20_64</b>	-1.809	0.163752359
<b>ID_A00_B99</b>	-12.260	4.7375E-06
<b>ID_C00_D48</b>	0.9495	2.584417128
<b>ID_D50_D89</b>	-12.316	4.48084E-06
<b>ID_E00_E90</b>	-2.757	0.063488277
<b>ID_F00_F99</b>	-3.638	0.026318057
<b>ID_G00_G99</b>	12.291	217618.7592
<b>ID_I00_I99</b>	1.639	5.150016919
<b>ID_J00_J99</b>	-0.917	0.399716393
<b>ID_K00_K93</b>	-1.338	0.262291185
<b>ID_L00_L99</b>	-12.539	3.58555E-06
<b>ID_M00_M99</b>	-12.100	5.55896E-06
<b>ID_N00_N99</b>	-2.887	0.055732044
<b>ID_O00_O99</b>	-10.611	2.46484E-05
<b>ID_P00_P96</b>	-5.505	0.004067609
<b>ID_Q00_Q99</b>	-8.969	0.000127257
<b>ID_R00_R99</b>	-12.213	4.96449E-06



**Región: Aysén del Gral. C. Ibáñez del Campo**

Tabla 187: *Evaluación - Coeficientes del modelo de RLO de Aysen*

<b>Variables</b>	<b>Coef</b>	<b>S.E</b>	<b>Wald Z</b>	<b>Wald</b>	<b>Pr(&gt; z )</b>
<b>y&gt;=2</b>	-63.4667	13.3536	-4.75	22.56	<0.0001
<b>y&gt;=3</b>	-65.5899	13.3560	-4.91	24.11	<0.0001
<b>y&gt;=4</b>	-67.6441	13.3584	-5.06	25.60	<0.0001
<b>y&gt;=5</b>	-69.1538	13.3657	-5.17	26.73	<0.0001
<b>TYear</b>	0.0314	0.0067	4.73	22.37	<0.0001
<b>TMonth</b>	0.0035	0.0115	0.31	0.10	0.7603
<b>DLocationID</b>	-0.6568	0.0813	-8.08	65.29	<0.0001
<b>SexID</b>	-0.6058	0.0853	-7.10	50.41	<0.0001
<b>RAgeID_0_10</b>	-2.9230	0.7217	-4.05	16.40	<0.0001
<b>RAgeID_11_19</b>	-3.1111	0.5936	-5.24	27.46	<0.0001
<b>RAgeID_20_64</b>	-0.8257	0.0881	-9.38	87.98	<0.0001
<b>ID_A00_B99</b>	-11.5979	47.1153	-0.25	0.06	0.8056
<b>ID_C00_D48</b>	-0.0945	0.1193	-0.79	0.62	0.4284
<b>ID_D50_D89</b>	-11.6540	101.2676	-0.12	0.01	0.9084
<b>ID_E00_E90</b>	-3.3947	0.5913	-5.74	32.95	<0.0001
<b>ID_F00_F99</b>	-4.1201	1.0091	-4.08	16.65	<0.0001
<b>ID_G00_G99</b>	-11.8220	44.6055	-0.27	0.07	0.7910
<b>ID_I00_I99</b>	-0.0143	0.1204	-0.12	0.01	0.9057
<b>ID_J00_J99</b>	-1.1373	0.1660	-6.85	46.92	<0.0001
<b>ID_K00_K93</b>	-1.8435	0.2413	-7.64	58.37	<0.0001
<b>ID_L00_L99</b>	-12.0576	103.7700	-0.12	0.01	0.9075
<b>ID_M00_M99</b>	-11.7214	81.8781	-0.14	0.02	0.8862
<b>ID_N00_N99</b>	-3.2759	0.5924	-5.53	30.58	<0.0001
<b>ID_O00_O99</b>	-10.5687	218.4183	-0.05	0.003	0.9614
<b>ID_P00_P96</b>	-8.8413	65.4153	-0.14	0.020	0.8925
<b>ID_Q00_Q99</b>	-9.8321	61.1632	-0.16	0.03	0.8723
<b>ID_R00_R99</b>	-2.8762	0.4263	-6.75	45.56	<0.0001

Tabla 188: *Evaluación - Probabilidades del modelo de RLO de Aysen*

<b>Variablen</b>	<b>B</b>	<b>EXP B</b>
<b>y&gt;=2</b>	-634.667	-
<b>y&gt;=3</b>	-655.899	-
<b>y&gt;=4</b>	-676.441	-
<b>y&gt;=5</b>	-691.538	-
<b>TYear</b>	0.0314	1.031898181
<b>TMonth</b>	0.0035	1.003506132
<b>DLocationID</b>	-0.6568	0.518507908
<b>SexID</b>	-0.6058	0.545637742
<b>RAgeID_0_10</b>	-2.923	5.37721E-02
<b>RAgeID_11_19</b>	-3.111	0.044551921
<b>RAgeID_20_64</b>	-0.8257	0.437928335
<b>ID_A00_B99</b>	-11.598	9.18536E-06
<b>ID_C00_D48</b>	-0.0945	0.909827735
<b>ID_D50_D89</b>	-11.654	8.68425E-06
<b>ID_E00_E90</b>	-3.395	0.033550618
<b>ID_F00_F99</b>	-4.120	0.01624289
<b>ID_G00_G99</b>	-11.822	7.34126E-06
<b>ID_I00_I99</b>	-0.0143	0.985801759
<b>ID_J00_J99</b>	-1.137	0.3206837
<b>ID_K00_K93</b>	-1.844	0.158262537
<b>ID_L00_L99</b>	-12.058	5.80031E-06
<b>ID_M00_M99</b>	-11.721	8.11822E-06
<b>ID_N00_N99</b>	-3.276	0.037782849
<b>ID_O00_O99</b>	-10.569	2.57082E-05
<b>ID_P00_P96</b>	-8.841	0.000144635
<b>ID_Q00_Q99</b>	-9.832	5.36999E-05
<b>ID_R00_R99</b>	-2.8762	0.056348481

## Series de tiempo con el modelo ARIMA

Los resultados de cada una de las regiones de Chile se detallan a continuación utilizando el modelo ARIMA para pronosticar la mortalidad.

### Región de Santiago

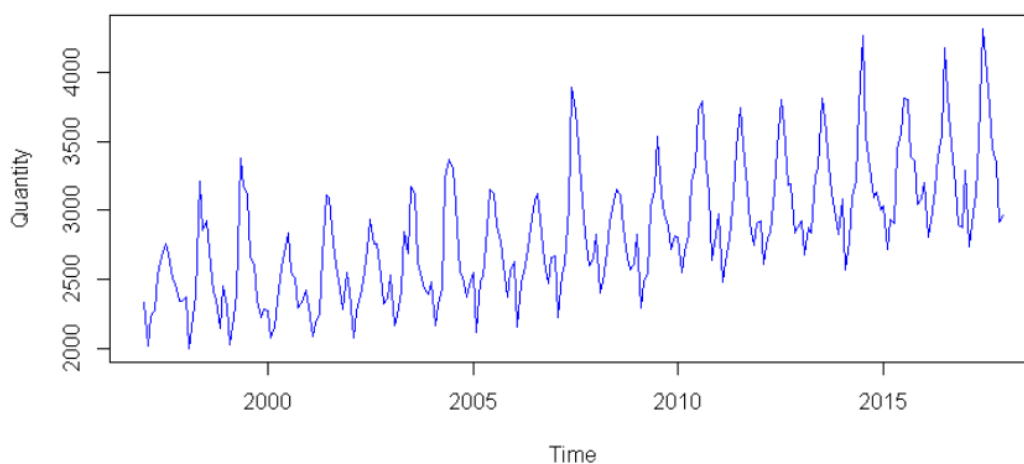


Figura 64: Evaluación - Comportamiento de la mortalidad de Santiago

Tabla 189: Evaluación - Resumen del modelo ARIMA de Santiago

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Santiago	-5.1146	141.4304	98.4152	-0.5057	3.3578	0.6693

En la prueba de Ljung-Box el modelo presenta un buen ajuste, debido a que el valor es de 0.8549, lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

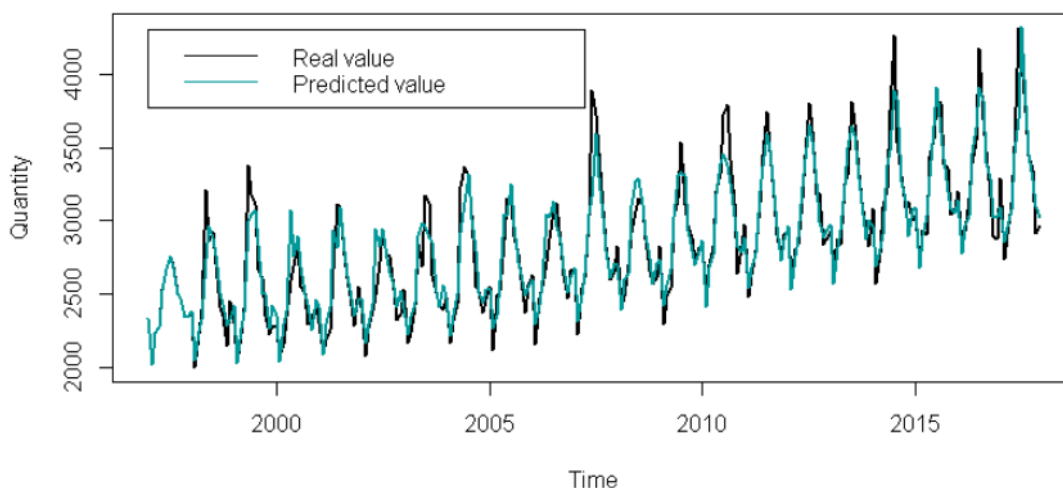


Figura 65: Evaluación - Ajuste del modelo ARIMA de Santiago

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Santiago

Tabla 190: Evaluación - Predicciones estimadas para la región de Santiago

Month	Point Forecast	LO 95	HI 95
Jan 2018	3193.842	2907.998	3479.687
Feb 2018	2816.289	2504.709	3127.868
Mar 2018	3009.854	2693.666	3326.042
Apr 2018	3153.953	2836.905	3471.001
May 2018	3505.101	3187.891	3822.310
Jun 2018	3838.065	3520.825	4155.305
Jul 2018	4099.691	3782.446	4416.937
Aug 2018	3808.616	3491.369	4125.863
Sep 2018	3512.357	3195.110	3829.604
Oct 2018	3349.532	3032.285	3666.779
Nov 2018	3096.537	2779.290	3413.784
Dec 2018	3119.987	2802.740	3437.233

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

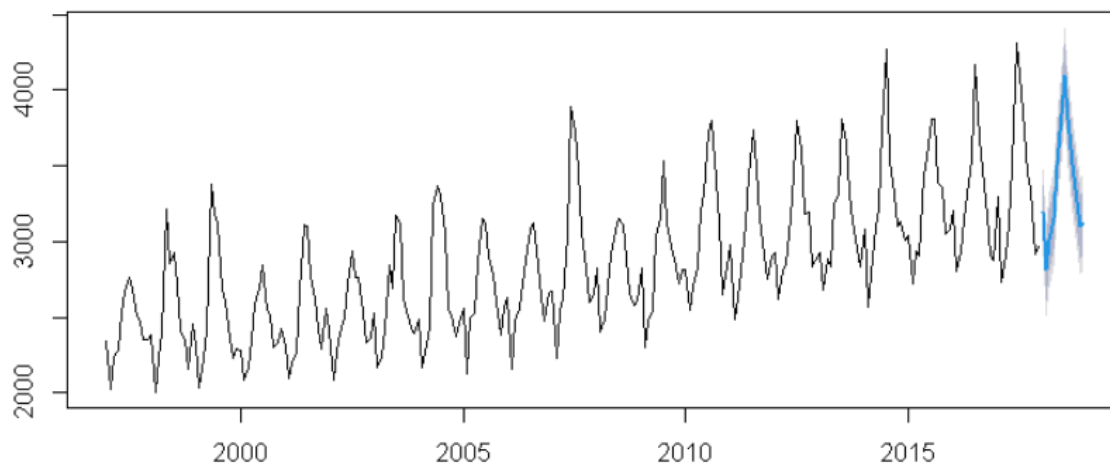


Figura 66: Evaluación - Pronóstico de la mortalidad de Santiago

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Valparaíso

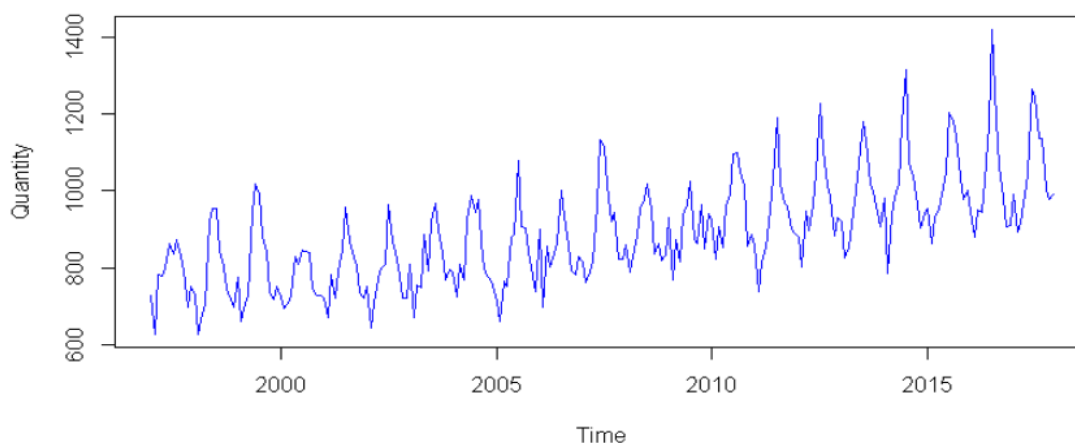


Figura 67: Evaluación - Comportamiento de la mortalidad de Valparaíso

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 191: Evaluación - Resumen del modelo ARIMA de Valparaíso

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Valparaíso	-1.8786	47.48408	35.6638	-0.5754	3.9762	0.6722

En la prueba Box-Ljung el modelo presenta un buen ajuste, debido a que el valor es de 0.8549 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

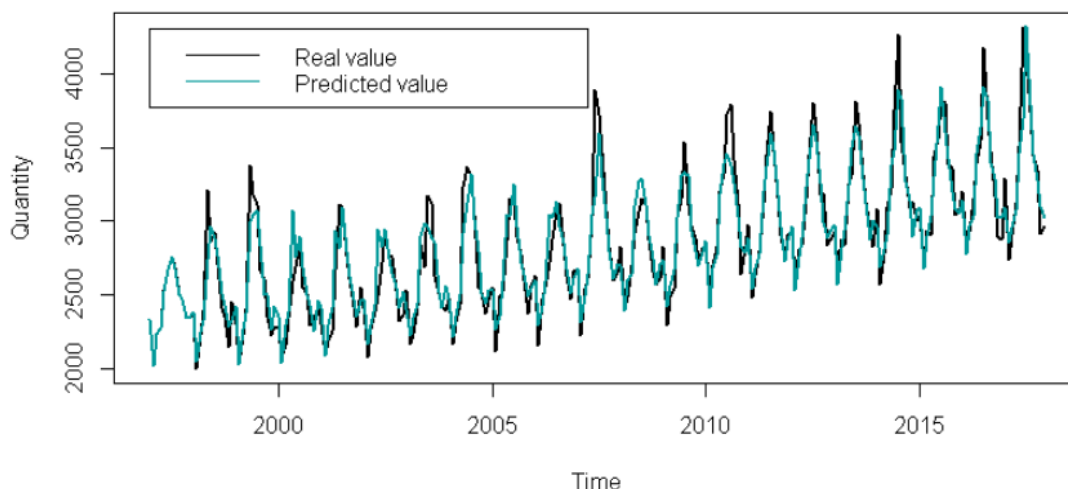


Figura 68: Evaluación - Ajuste del modelo ARIMA de Valparaíso

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Santiago

Tabla 192: Evaluación - Predicciones estimadas para la región de Valparaíso

Month	Point Forecast	LO 95	HI 95
Jan 2018	1002.2955	905.9209	1098.670
Feb 2018	895.2250	793.0950	997.355
Mar 2018	971.6100	868.2945	1074.926
Apr 2018	989.2782	885.9627	1092.594
May 2018	1057.4947	954.1792	1160.810
Jun 2018	1173.0317	1069.7162	1276.347
Jul 2018	1311.2163	1207.9008	1414.532
Aug 2018	1177.1204	1073.8049	1280.436
Sep 2018	1112.1830	1008.8675	1215.499
Oct 2018	1043.8121	940.4966	1147.128
Nov 2018	977.8208	874.5053	1081.136
Dec 2018	992.1644	888.8489	1095.480

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

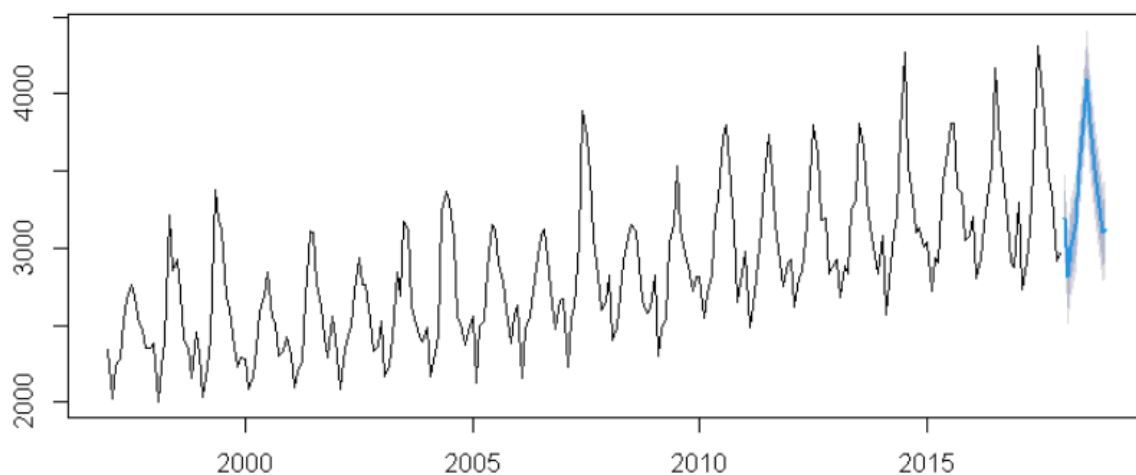


Figura 69: Evaluación - Pronóstico de la mortalidad de Valparaíso

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Biobío

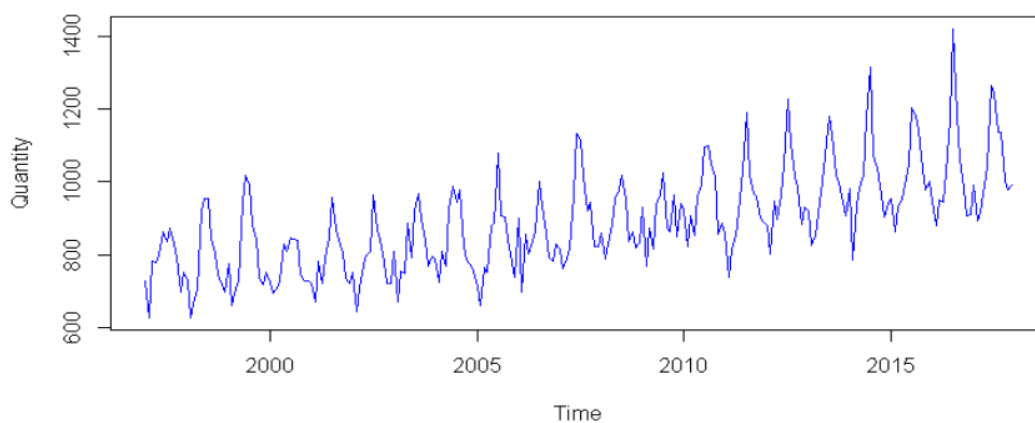


Figura 70: Evaluación - Comportamiento de la mortalidad de Biobío

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 193: Evaluación - Resumen del modelo ARIMA de Biobío

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Biobío	-1.1608	41.5533	30.2635	-0.5345	4.3288	0.6357

En la prueba de Ljung-Box el modelo presenta un buen ajuste, debido a que el valor es de 0.6607 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

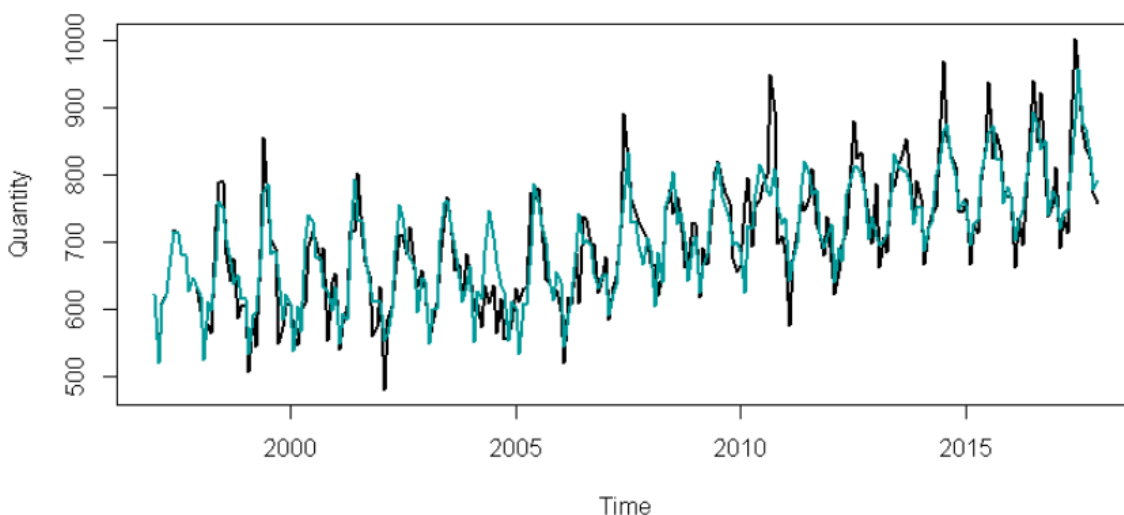


Figura 71: Evaluación - Ajuste del modelo ARIMA de Biobío

Nota: Este gráfico muestra los datos pronosticados o reales para la región Biobío.

Tabla 194: Evaluación - Predicciones estimadas para la región de Biobío

Month	Point Forecast	LO 95	HI 95
Jan 2018	777.3424	692.7832	861.9017
Feb 2018	710.2703	617.9131	802.6275
Mar 2018	761.3492	668.5039	854.1945
Apr 2018	756.2779	662.9772	849.5785
May 2018	822.7330	729.0072	916.4587
Jun 2018	869.9873	775.8646	964.1099
Jul 2018	918.8730	824.3797	1013.3663
Aug 2018	870.5122	775.6726	965.3518
Sep 2018	878.7548	783.5916	973.9180
Oct 2018	835.9764	740.5107	931.4420
Nov 2018	776.3903	680.6419	872.1387
Dec 2018	791.7017	695.6903	887.7132

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%



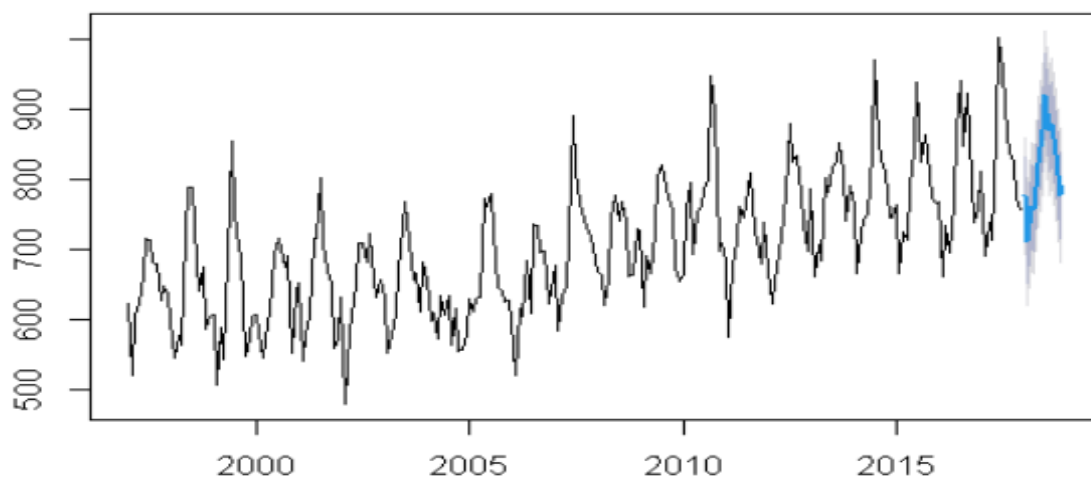


Figura 72: Evaluación - Pronóstico de la mortalidad de Biobío

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Maule

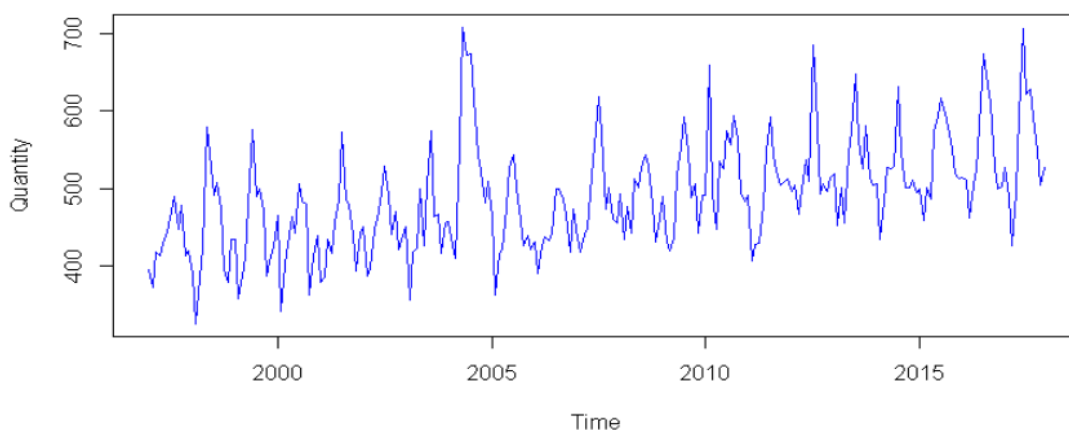


Figura 73: Evaluación - Comportamiento de la mortalidad de Maule

Nota: Este gráfico muestra el comportamiento de las muertes en la región del Maule.

Tabla 195: Evaluación - Resumen del modelo ARIMA de Maule

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Maule	8.8537	40.2790	28.8110	1.2864	5.7546	0.7015

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.0820 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

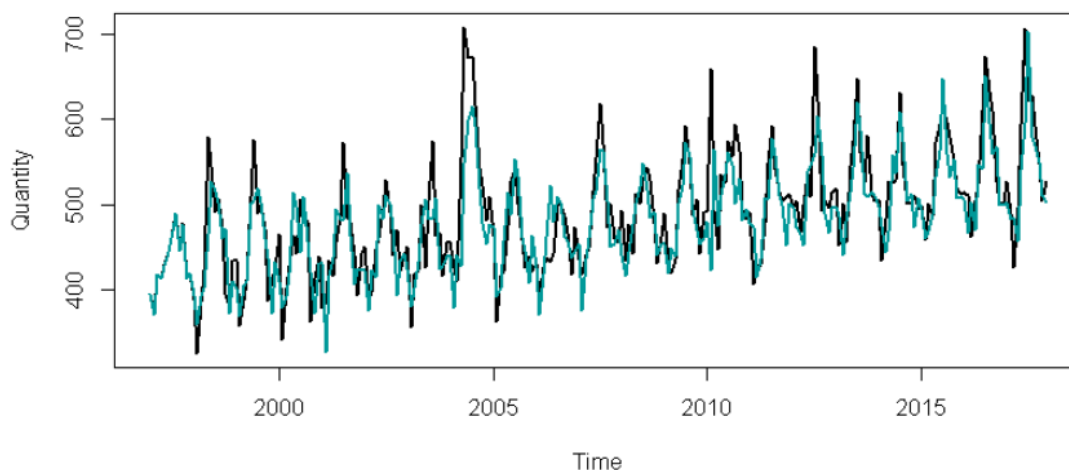


Figura 74: Evaluación - Ajuste del modelo ARIMA de Maule

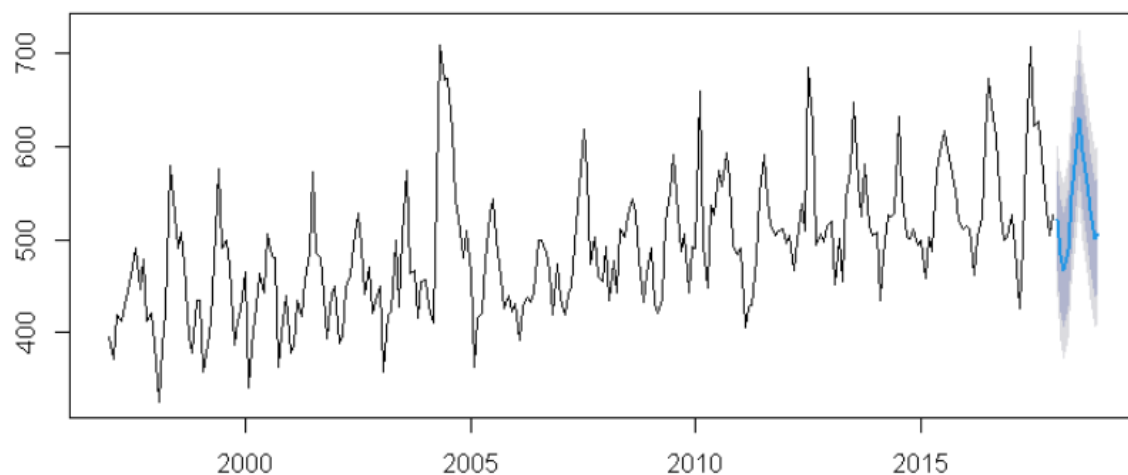
Nota: Este gráfico muestra los datos pronosticados o reales para la región del Maule.

Tabla 196: Evaluación - Predicciones estimadas para la región de Maule

Month	Point Forecast	LO 95	HI 95
Jan 2018	520.7901	4392.123	602.3680
Feb 2018	483.2253	3912.429	575.2077
Mar 2018	466.6171	372.0086	561.2256
Apr 2018	489.1514	393.8428	584.4600
May 2018	547.0002	451.5025	642.4979
Jun 2018	597.0977	501.5487	692.6466
Jul 2018	631.2951	535.7322	726.8579
Aug 2018	595.3417	499.7751	690.9083
Sep 2018	559.9763	464.4086	655.5439
Oct 2018	531.1427	435.5748	626.7106
Nov 2018	500.5183	404.9503	596.0863
Dec 2018	505.7375	410.1695	601.3055

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

Figura 75: Evaluación - Pronóstico de la mortalidad de Maule



Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Araucanía

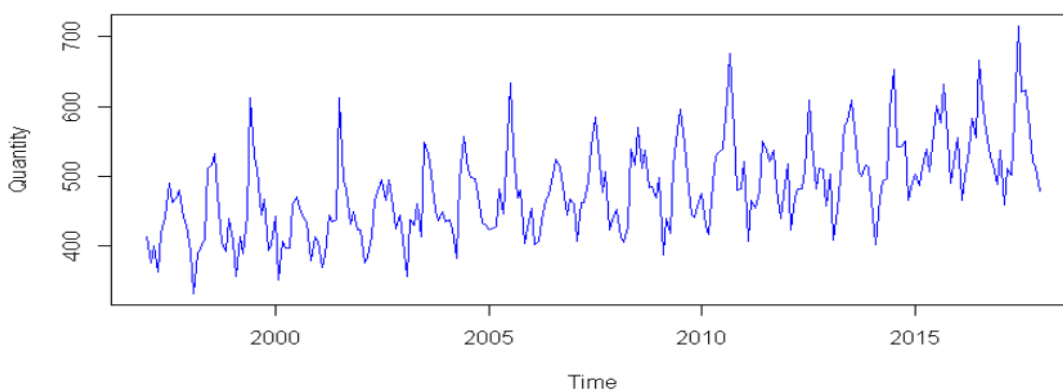


Figura 76: Evaluación - Comportamiento de la mortalidad de Araucanía

Nota: Este gráfico muestra el comportamiento de las muertes en la región de la Araucanía.

Tabla 197: Evaluación - Resumen del modelo ARIMA de Araucanía

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Araucanía	-1.5589	30.2382	22.1755	-0.7386	4.5570	0.6203

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.9435 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

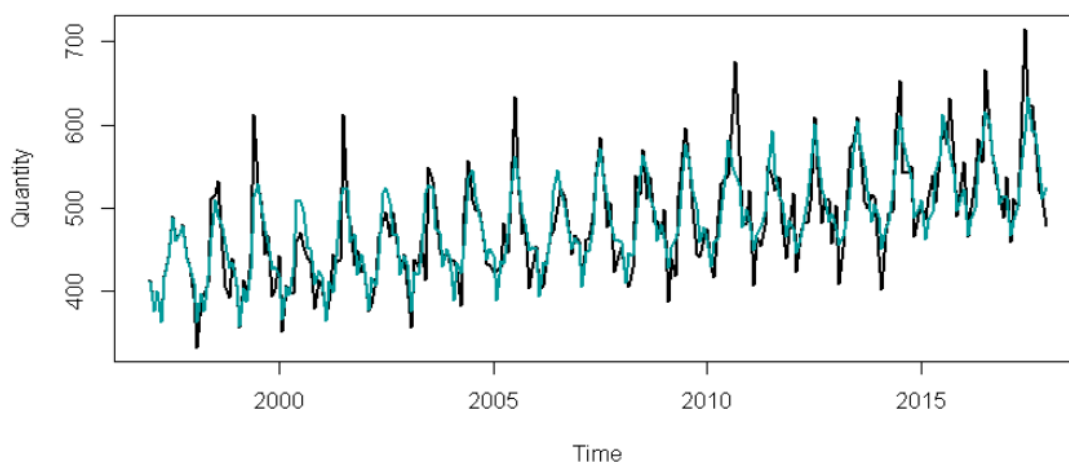


Figura 77: Evaluación - Ajuste del modelo ARIMA de Araucanía

Nota: Este gráfico muestra los datos pronosticados o reales para la región de la Araucanía.

Tabla 198: Evaluación - Predicciones estimadas para la región de Araucanía

Month	Point Forecast	LO 95	HI 95
Jan 2018	517.0866	455.2995	578.8737
Feb 2018	469.7280	407.3361	532.1199
Mar 2018	490.5047	428.0966	552.9128
Apr 2018	508.2574	445.6336	570.8813
May 2018	548.1467	485.4452	610.8481
Jun 2018	572.2947	509.5653	635.0241
Jul 2018	626.3252	563.5857	689.0646
Aug 2018	594.1569	531.4138	656.9000
Sep 2018	579.4196	516.6752	642.1640
Oct 2018	555.0931	492.3482	617.8380
Nov 2018	515.1442	452.3991	577.8892
Dec 2018	529.1821	466.4370	591.9272

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

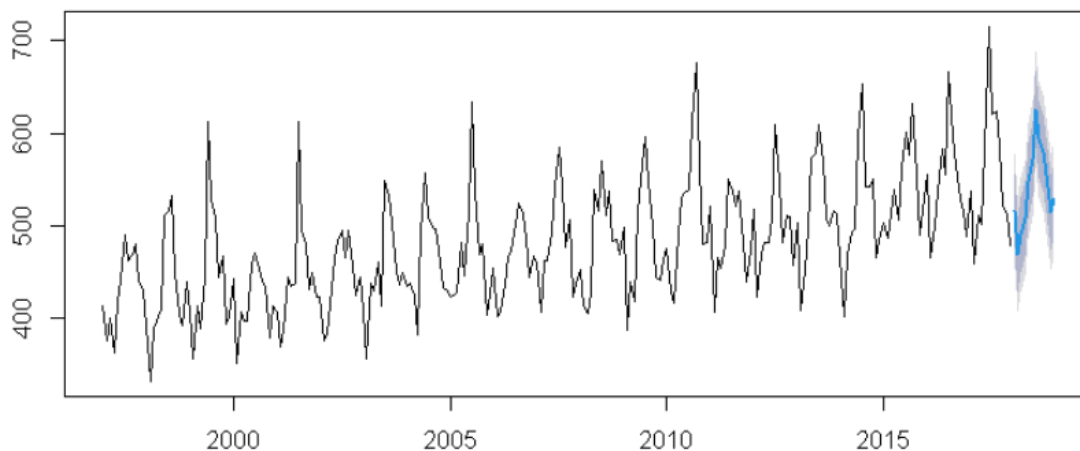


Figura 78: Evaluación - Pronóstico de la mortalidad de Araucanía

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Libertador General Bernardo O'Higgins

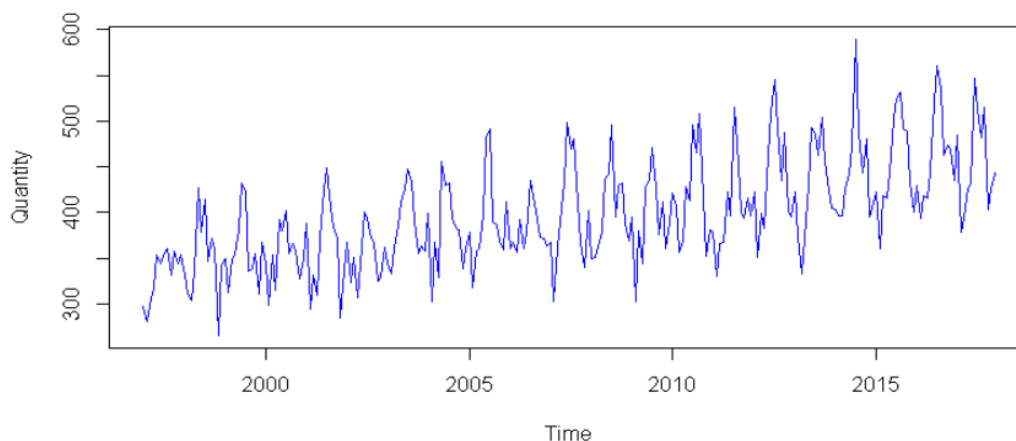


Figura 79: Evaluación - Comportamiento de la mortalidad de O'Higgins

Nota: Este gráfico muestra el comportamiento de las muertes en la región de O'Higgins.

Tabla 199: Evaluación - Resumen del modelo ARIMA de O'Higgins

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
O'Higgins	0.2544	26.9128	20.752	-0.4278	5.2084	0.6765

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.1396 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

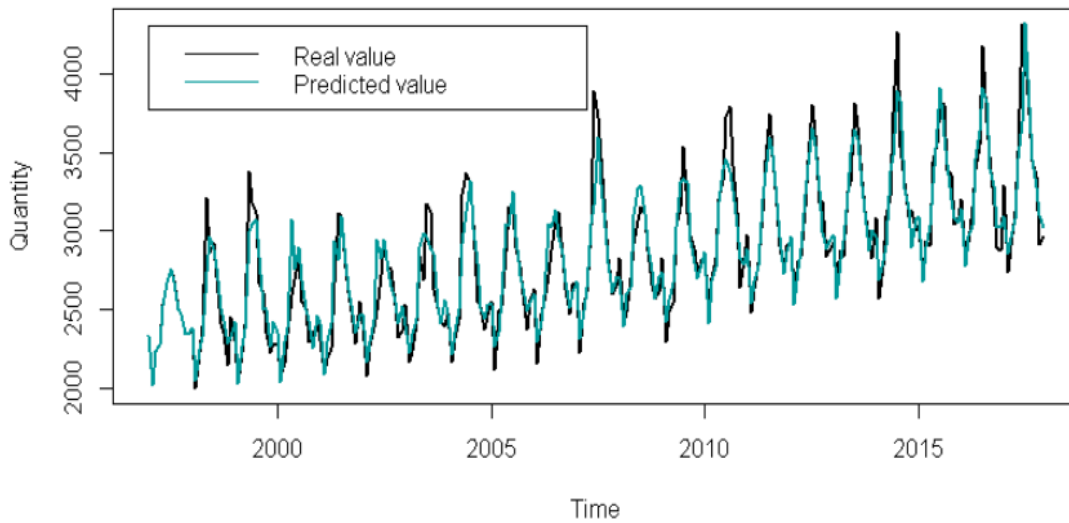


Figura 80: Evaluación - Ajuste del modelo ARIMA de O'Higgins

Nota: Este gráfico muestra los datos pronosticados o reales para la región O'Higgins.

Tabla 200: Evaluación - Predicciones estimadas para la región de O'Higgins

Month	Point Forecast	LO 95	HI 95
Jan 2018	441.8902	387.3833	496.3971
Feb 2018	398.4188	343.9119	452.9257
Mar 2018	420.7555	366.2486	475.2624
Apr 2018	428.5783	374.0714	483.0852
May 2018	466.7722	412.2653	521.2791
Jun 2018	514.0027	459.4958	568.5096
Jul 2018	561.6347	507.1278	616.1416
Aug 2018	526.4490	471.9421	580.9559
Sep 2018	489.6342	435.1273	544.1411
Oct 2018	492.2023	437.6954	546.7092
Nov 2018	447.7322	393.2252	502.2391
Dec 2018	434.5401	380.0332	489.0470

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

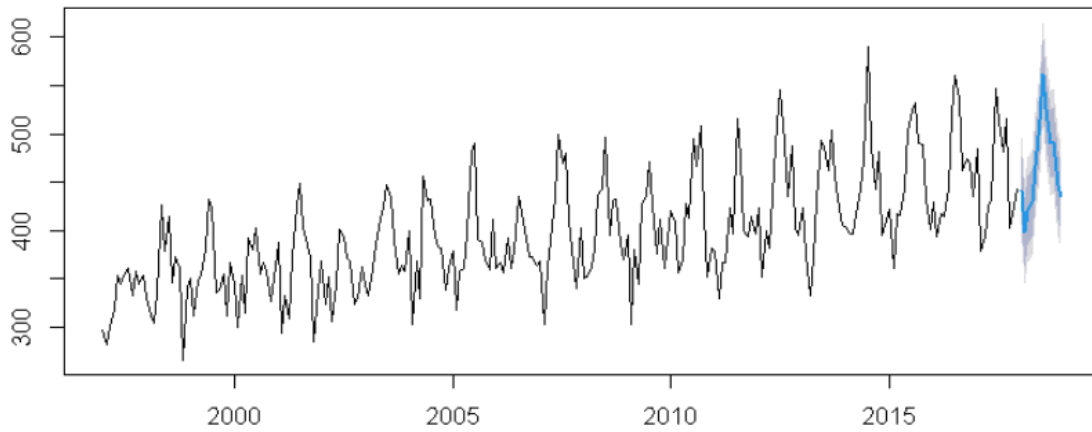


Figura 81: Evaluación - Pronóstico de la mortalidad de O'Higgins

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Los Lagos

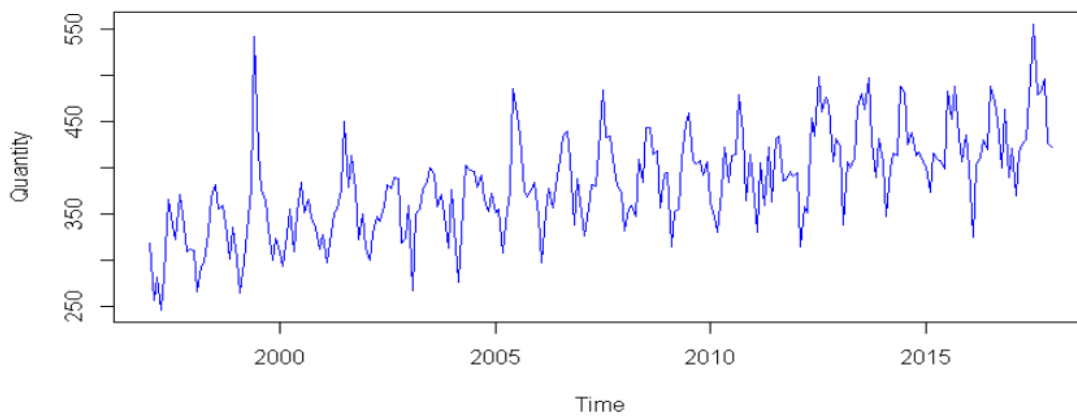


Figura 82: Evaluación - Comportamiento de la mortalidad de Los Lagos

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 201: Evaluación - Resumen del modelo ARIMA de Los Lagos

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Los Lagos	9.4401	29.9794	22.8530	2.09868	5.852786	0.81363

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.0118 lo cual es menor a 0.05 por lo que no existe ruido.

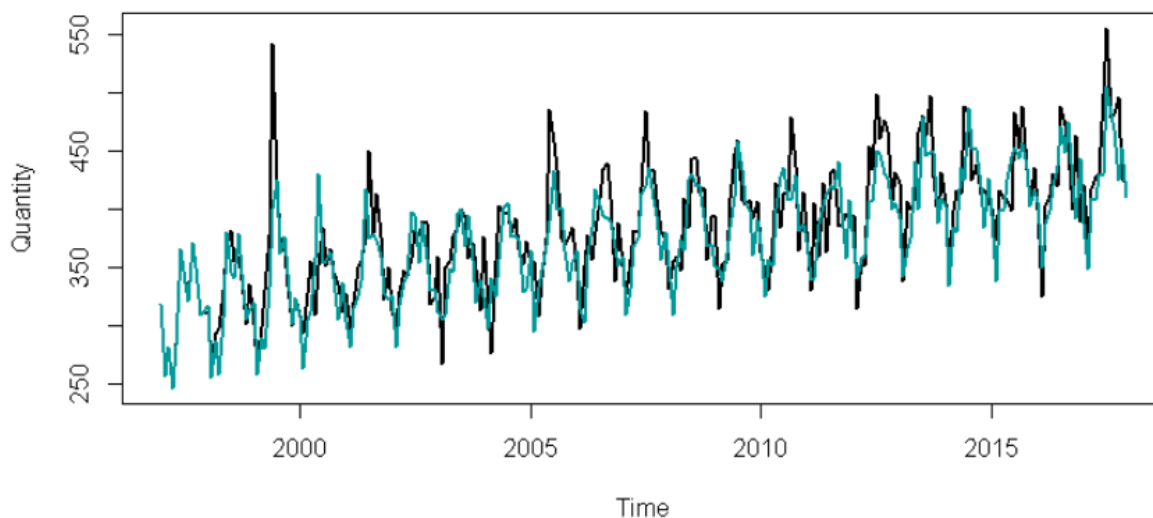


Figura 83: Evaluación - Ajuste del modelo ARIMA de Los Lagos

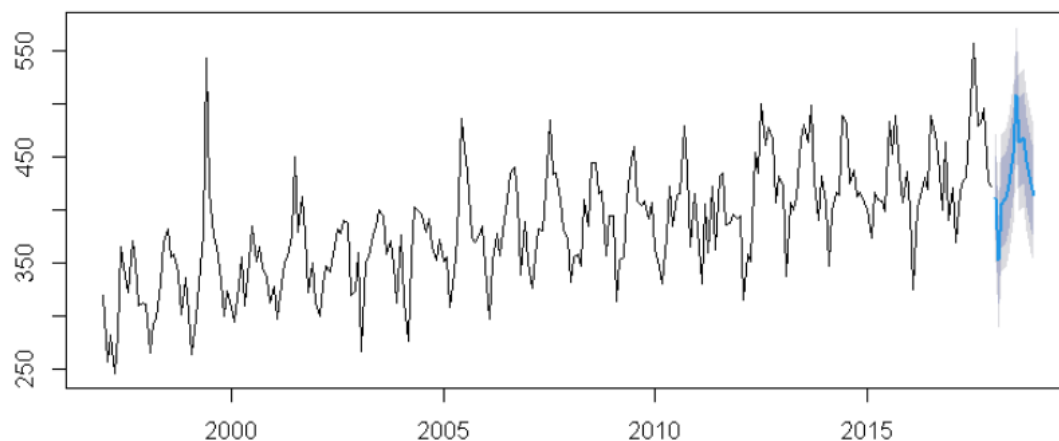
Nota: Este gráfico muestra los datos pronosticados o reales para la región de Santiago.

Tabla 202: Evaluación - Predicciones estimadas para la región de Los Lagos

Month	Point Forecast	LO 80	HI 80
Jan 2018	411.4199	350.9577	471.8821
Feb 2018	352.9695	288.6595	417.2794
Mar 2018	405.9243	341.1260	470.7226
Apr 2018	411.3150	346.4528	476.1771
May 2018	423.7277	358.8572	488.5983
Jun 2018	452.9678	388.0962	517.8395
Jul 2018	507.2693	442.3975	572.1411
Aug 2018	463.0276	398.1557	527.8994
Sep 2018	467.8515	402.9797	532.7233
Oct 2018	445.3357	380.4639	510.2075
Nov 2018	424.5901	359.7182	489.4619
Dec 2018	414.4443	349.5725	479.3161



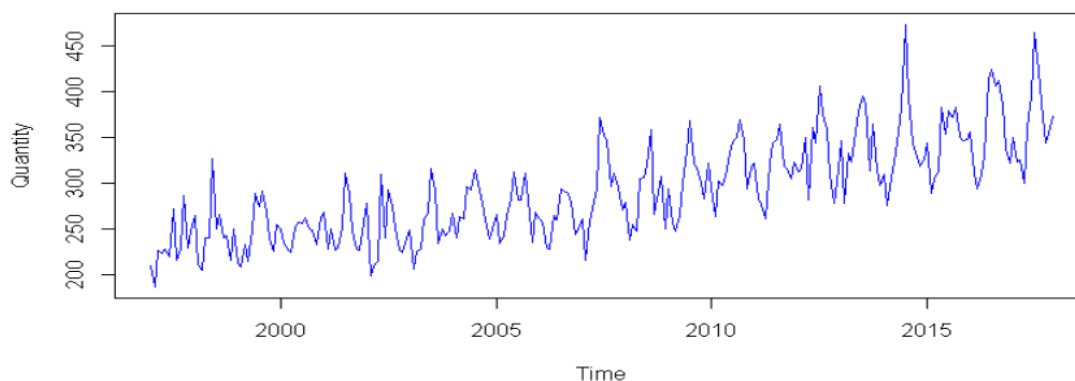
*Nota:* La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.



*Figura 84:* Evaluación - Pronóstico de la mortalidad de Los Lagos

*Nota:* Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Coquimbo



*Figura 85:* Evaluación - Comportamiento de la mortalidad de Coquimbo

*Nota:* Este gráfico muestra el comportamiento de las muertes en la región de Coquimbo.

*Tabla 203:* Evaluación - Resumen del modelo ARIMA de Coquimbo

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Coquimbo	-1.1392	21.87694	16.5561	-1.1040	5.7282	0.6961

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.6354 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

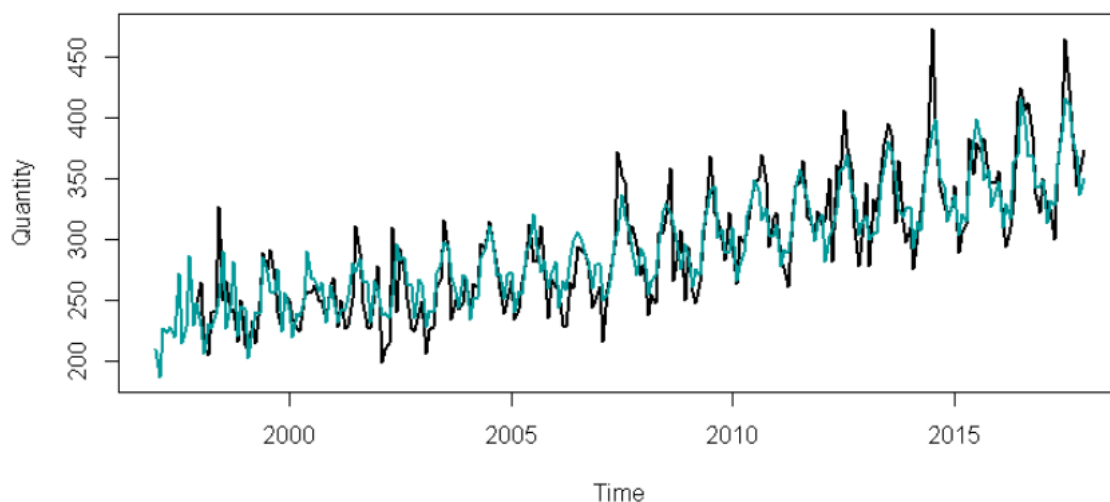


Figura 86: Evaluación - Ajuste del modelo ARIMA de Coquimbo

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Santiago

Tabla 204: Evaluación - Predicciones estimadas para la región de Coquimbo

Month	Point Forecast	LO 95	HI 95
Jan 2018	365.8384	321.4167	410.2601
Feb 2018	324.3178	278.6068	370.0288
Mar 2018	335.9570	290.1712	381.7428
Apr 2018	331.5312	285.7410	377.3215
May 2018	373.3729	327.5824	419.1633
Jun 2018	397.7341	351.9436	443.5246
Jul 2018	428.9499	383.1594	474.7404
Aug 2018	409.5415	363.7510	455.3320
Sep 2018	386.0602	340.2697	431.8507
Oct 2018	370.6407	324.8502	416.4312
Nov 2018	350.5337	304.7432	396.3242
Dec 2018	356.9016	311.1114	402.6918

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

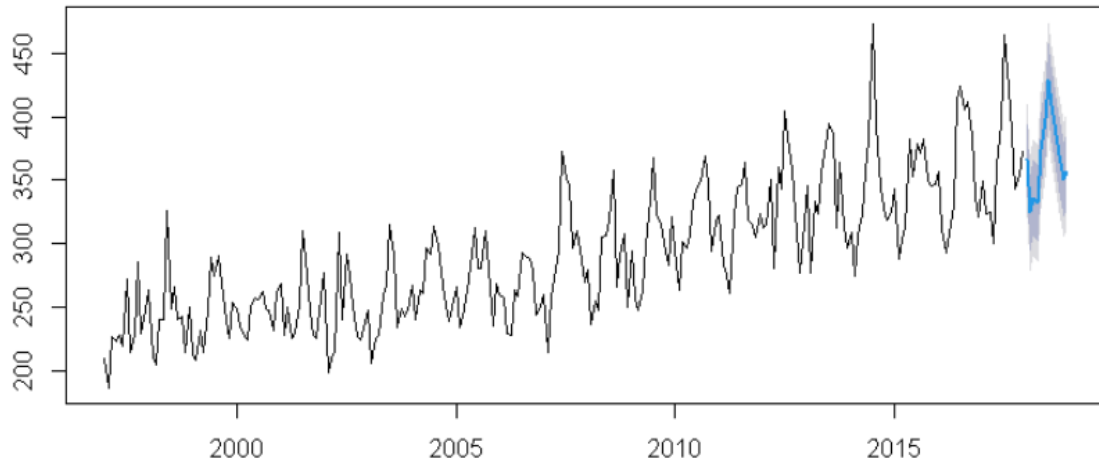


Figura 87: Evaluación - Pronóstico de la mortalidad de Coquimbo

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Ñuble

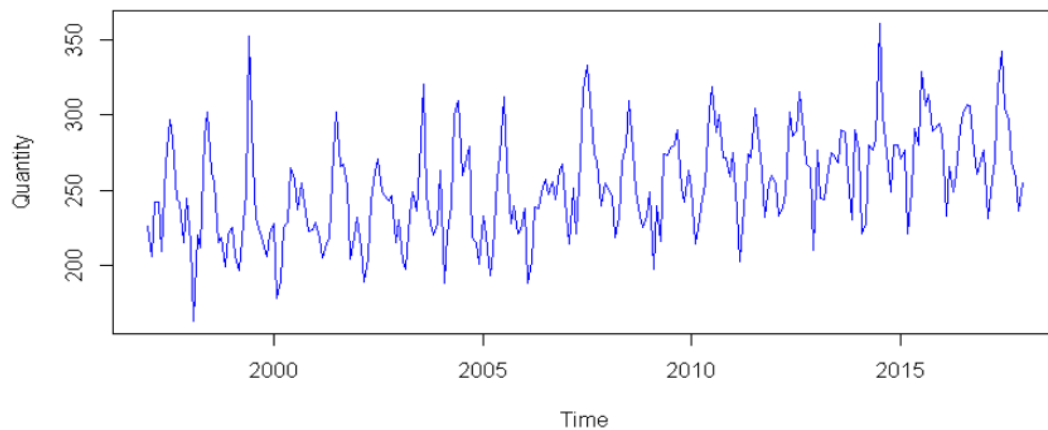


Figura 88: Evaluación - Comportamiento de la mortalidad de Ñuble

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 205: Evaluación - Resumen del modelo ARIMA de Ñuble

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Ñuble	-1.2505	19.2729	14.5753	-1.0451	5.7922	0.6446

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.8921 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

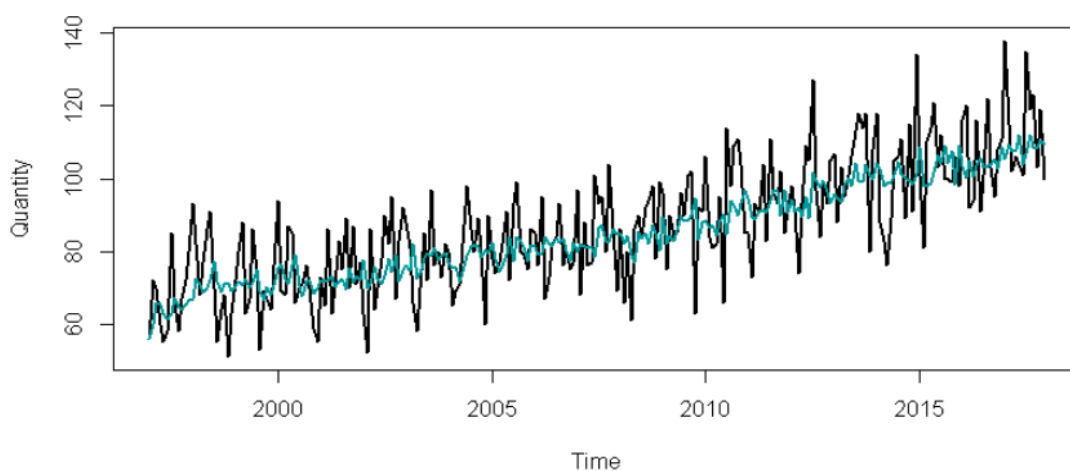


Figura 89: Evaluación - Ajuste del modelo ARIMA de Ñuble

Nota: Este gráfico muestra los datos pronosticados o reales para la región Ñuble

Tabla 206: Evaluación - Predicciones estimadas para la región de Ñuble

Month	Point Forecast	LO 95	HI 95
Jan 2018	274.1121	234.8157	313.4084
Feb 2018	247.2777	207.3215	287.2340
Mar 2018	242.5323	202.5171	282.5475
Apr 2018	263.0614	223.0269	303.0958
May 2018	292.4865	252.4405	332.5325
Jun 2018	301.9953	261.9413	342.0492
Jul 2018	324.2991	284.2394	364.3588
Aug 2018	309.2590	269.1952	349.3227
Sep 2018	297.3910	257.3243	337.4577
Oct 2018	275.3429	235.2740	315.4117
Nov 2018	273.5179	233.4475	313.5883
Dec 2018	277.6878	237.6164	317.7591

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

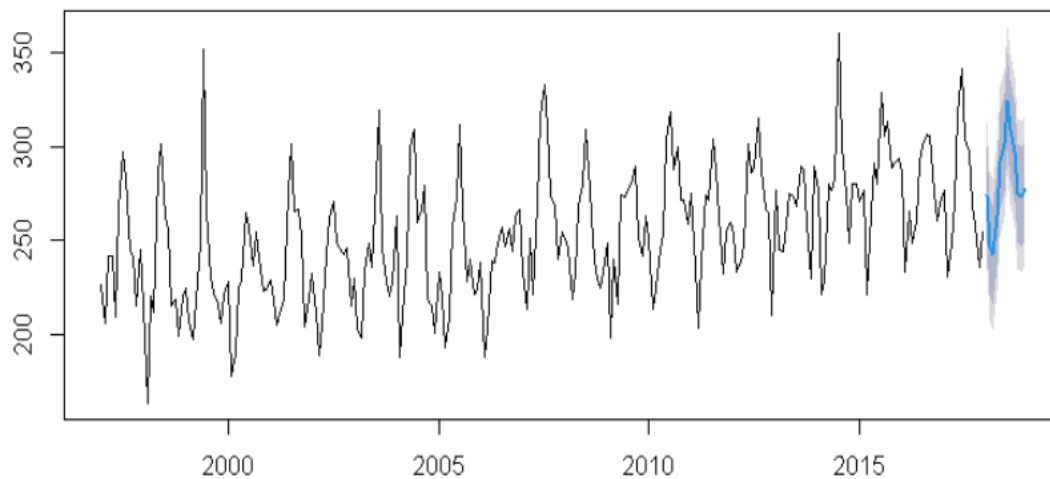


Figura 90: Evaluación - Pronóstico de la mortalidad de Ñuble

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Antofagasta

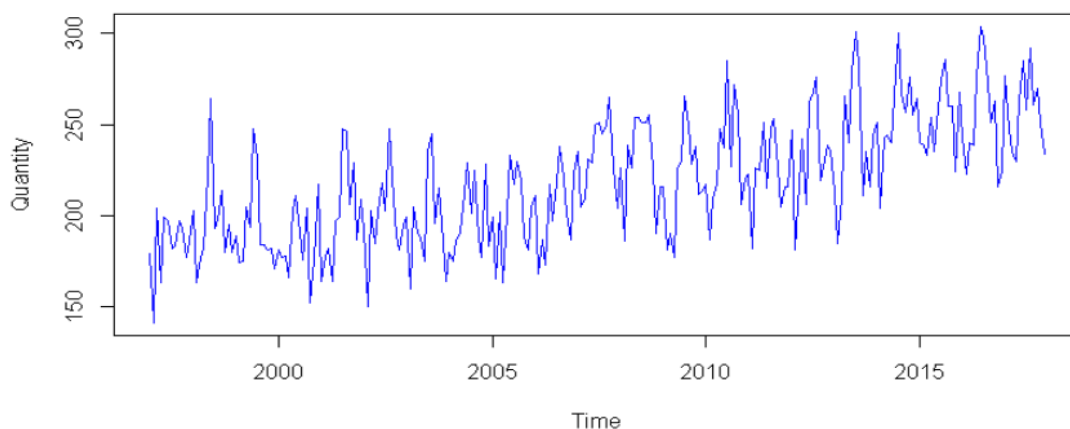


Figura 91: Evaluación - Comportamiento de la mortalidad de Antofagasta

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 207: Evaluación - Resumen del modelo ARIMA de Antofagasta

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Antofagasta	-0.1138	23.0759	18.6217	-0.8233	8.6923	0.9454

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.0326 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

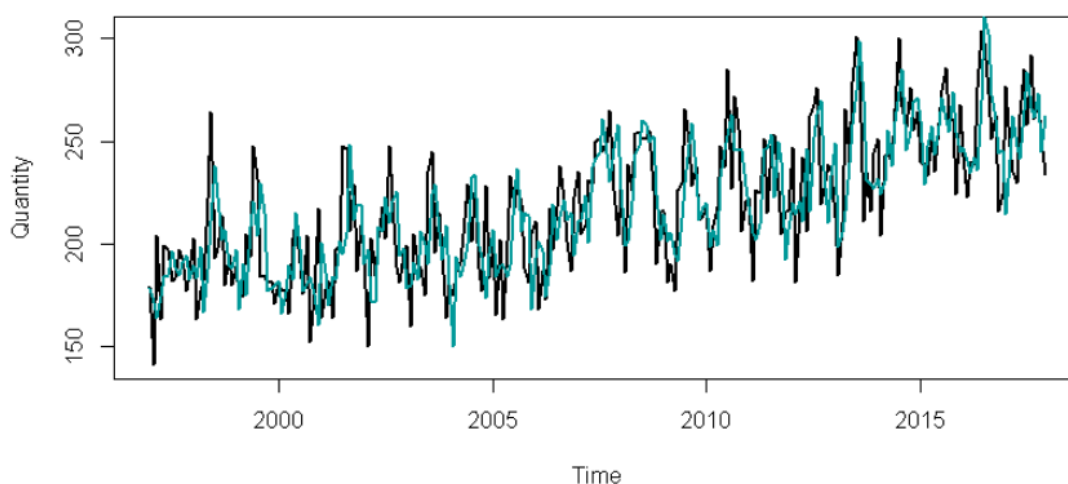


Figura 92: Evaluación - Ajuste del modelo ARIMA de Antofagasta

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Antofagasta

Tabla 208: Evaluación - Predicciones estimadas para la región de Antofagasta

Month	Point Forecast	LO 95	HI 95
Jan 2018	257.7984	212.1148	303.4819
Feb 2018	238.3418	186.3221	290.3615
Mar 2018	241.4324	179.0928	303.7720
Apr 2018	238.5704	169.3704	307.7704
May 2018	257.9331	181.6906	334.1756
Jun 2018	271.0228	188.6775	353.3681
Jul 2018	261.7910	173.6202	349.9618
Aug 2018	266.2069	172.6346	359.7792
Sep 2018	251.9096	153.2042	350.6150
Oct 2018	257.5854	154.0126	361.1581
Nov 2018	240.9788	132.7523	349.2054
Dec 2018	237.9212	125.2352	350.6073

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

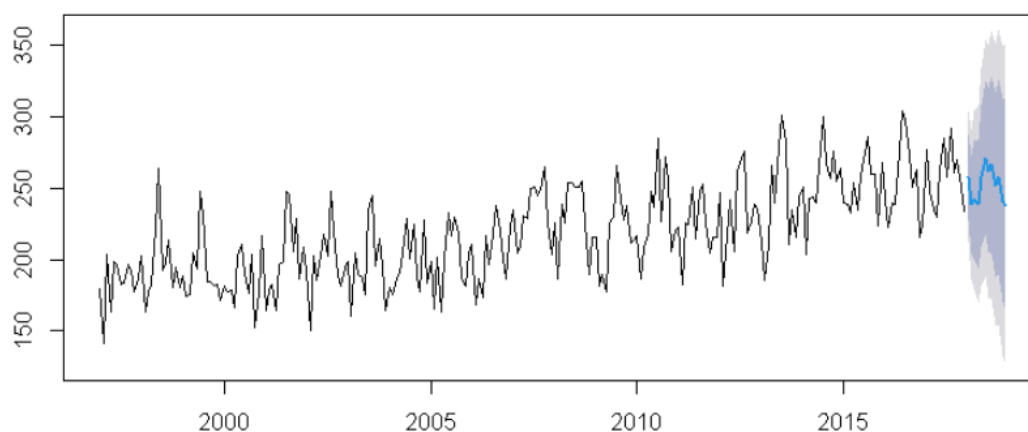


Figura 93: Evaluación - Pronóstico de la mortalidad de Antofagasta

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Los Ríos

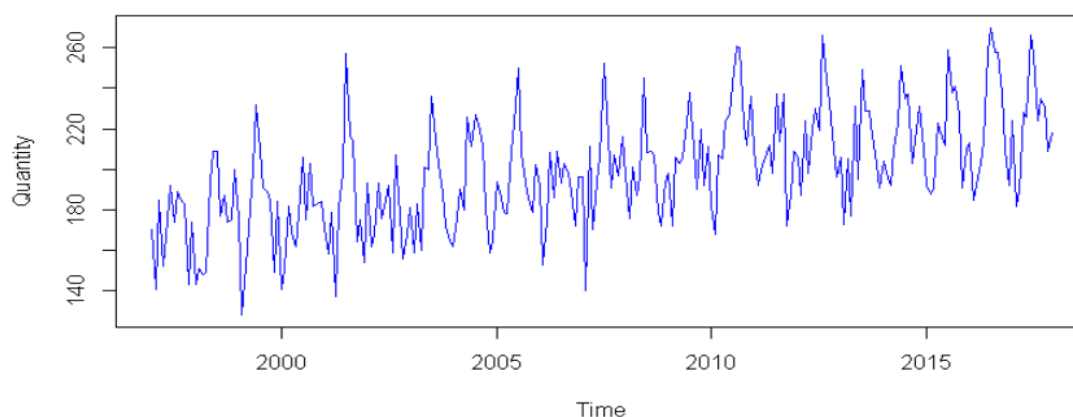


Figura 94: Evaluación - Comportamiento de la mortalidad de Los Ríos

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Los Ríos.

Tabla 209: Evaluación - Resumen del modelo ARIMA de Los Ríos

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Los Ríos	2.1141	17.8165	13.8879	0.2667	7.0812	0.7498

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.8122 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

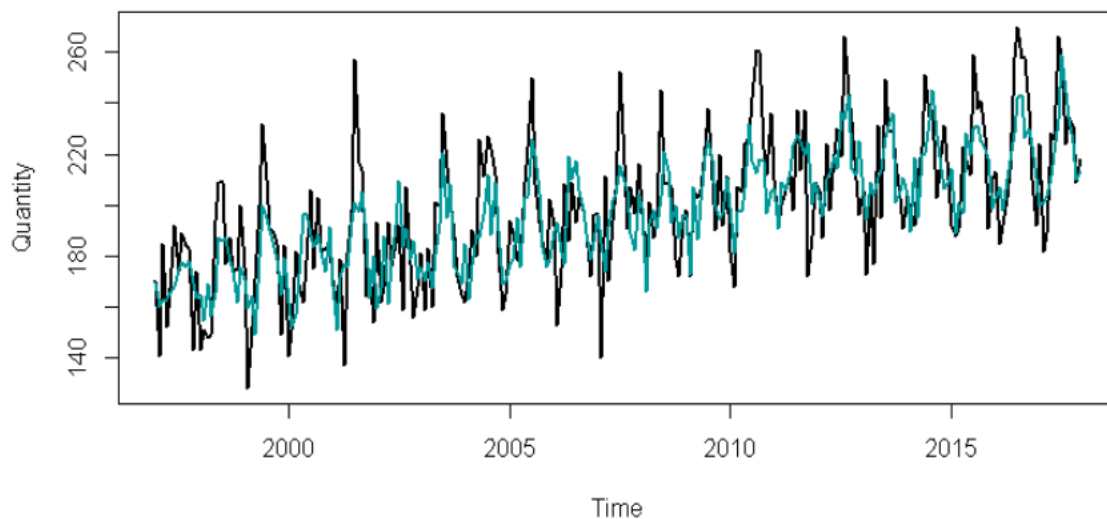


Figura 95: Evaluación - Ajuste del modelo ARIMA de Los Ríos

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Los Ríos.

Tabla 210: Evaluación - Predicciones estimadas para la región de Los Ríos

Month	Point Forecast	LO 95	HI 95
Jan 2018	222.0299	186.4693	257.5905
Feb 2018	200.4168	163.9315	236.9021
Mar 2018	203.9338	166.7863	241.0813
Apr 2018	218.1798	181.0271	255.3325
May 2018	222.1275	184.9701	259.2849
Jun 2018	249.6869	212.5295	286.8443
Jul 2018	252.2071	215.0471	289.3671
Aug 2018	239.1712	202.0036	276.3388
Sep 2018	241.9785	204.7984	279.1587
Oct 2018	232.6981	195.5012	269.8950
Nov 2018	216.4946	179.2779	253.7114
Dec 2018	211.6736	174.4347	248.9126



Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

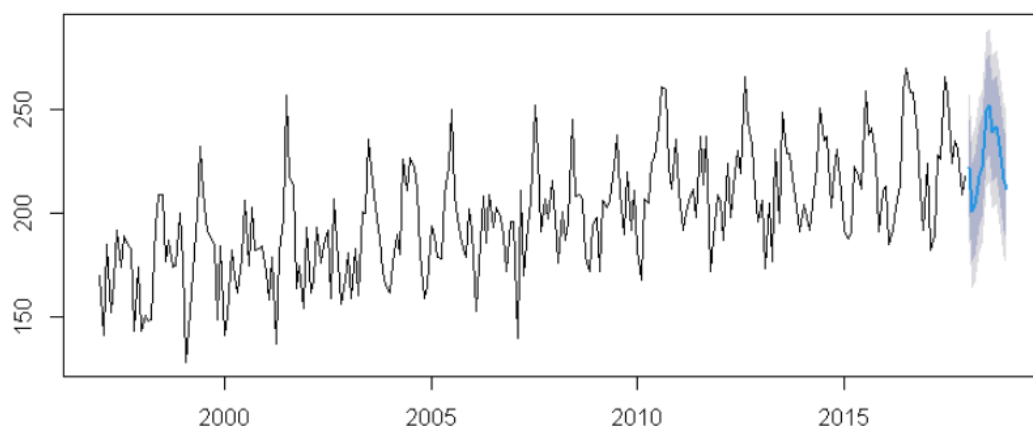


Figura 96: Evaluación - Pronóstico de la mortalidad de Los Ríos

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Atacama

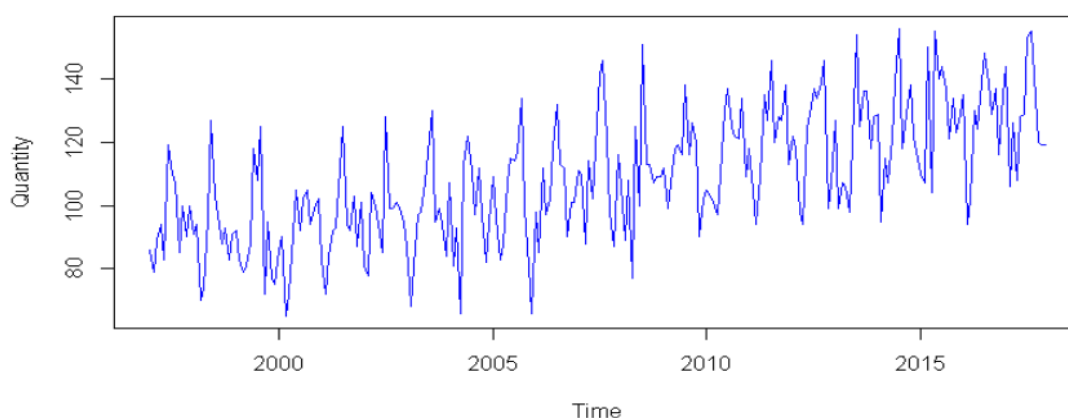


Figura 97: Evaluación - Comportamiento de la mortalidad de Atacama

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Valparaíso.

Tabla 211: Evaluación - Resumen del modelo ARIMA de Atacama

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Atacama	-1.4040	12.7431	9.7610	-0.0843	9.2153	0.8000

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.9957 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

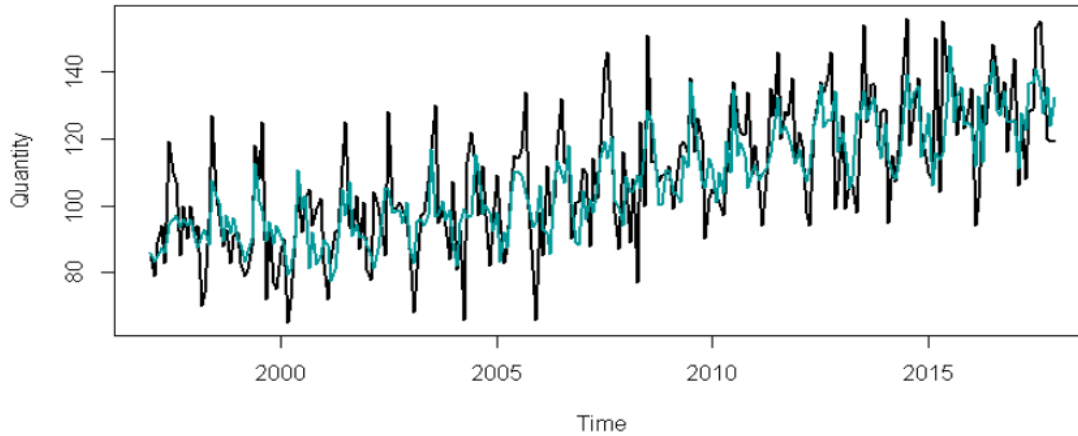


Figura 98: Evaluación - Ajuste del modelo ARIMA de Atacama

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Atacama.

Tabla 212: Evaluación - Predicciones estimadas para la región de Atacama

Month	Point Forecast	LO 95	HI 95
Jan 2018	136.7999	111.62313	161.9767
Feb 2018	110.2580	85.06359	135.4525
Mar 2018	119.8625	94.65040	145.0745
Apr 2018	122.5335	97.30384	147.7632
May 2018	127.6377	102.39036	152.8850
Jun 2018	132.8111	107.54617	158.0760
Jul 2018	144.1436	118.86104	169.4261
Aug 2018	141.9538	116.65369	167.2539
Sep 2018	133.8163	108.49863	159.1339
Oct 2018	129.0034	103.66820	154.3386
Nov 2018	121.9013	96.54854	147.2540
Dec 2018	128.0089	102.63858	153.3791

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

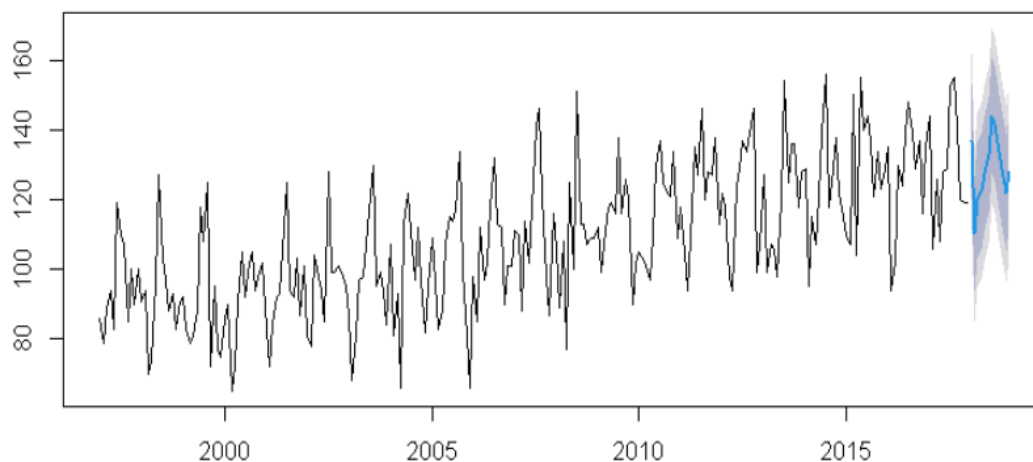


Figura 99: Evaluación - Pronóstico de la mortalidad de Atacama

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Tarapacá

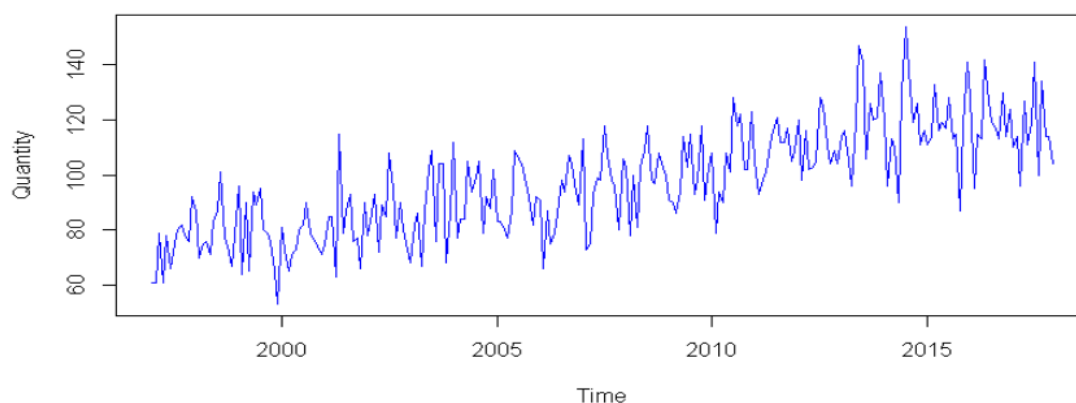


Figura 100: Evaluación - Comportamiento de la mortalidad de Tarapacá

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Tarapacá.

Tabla 213: Evaluación - Resumen del modelo ARIMA de Tarapacá

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Tarapacá	1.7380	11.8725	9.2250	0.4675	9.5728	0.7437

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.6607 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

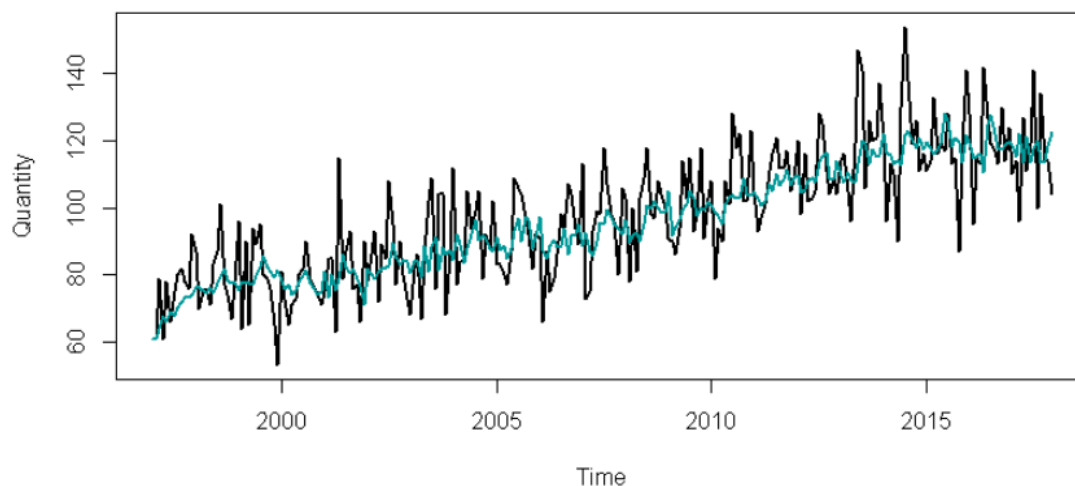


Figura 101: Evaluación - Ajuste del modelo ARIMA de Tarapacá

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Tarapacá.

Tabla 214: Evaluación - Predicciones estimadas para la región de Tarapacá

Month	Point Forecast	LO 95	HI 95
Jan 2018	116.8239	93.31985	140.3280
Feb 2018	112.4270	88.53312	136.3209
Mar 2018	114.1664	90.18991	138.1429
Apr 2018	116.4642	92.40535	140.5230
May 2018	123.0197	98.87885	147.1606
Jun 2018	118.5850	94.36243	142.8077
Jul 2018	116.9150	92.61089	141.2191
Aug 2018	114.3363	89.95095	138.7216
Sep 2018	116.9048	92.43853	141.3710
Oct 2018	119.8015	95.25453	144.3484
Nov 2018	115.9489	91.32153	140.5762
Dec 2018	116.5248	91.81737	141.2323

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%

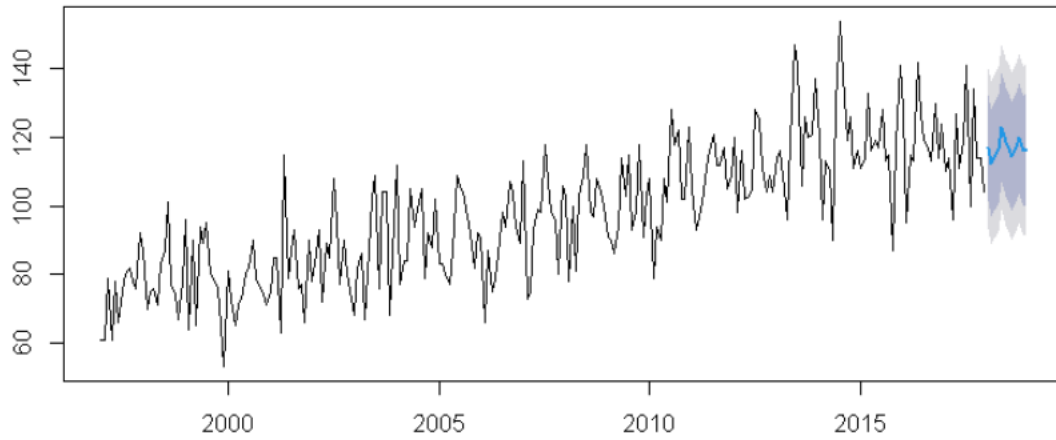


Figura 102: Evaluación - Pronóstico de la mortalidad de Tarapacá

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región Arica y Parinacota

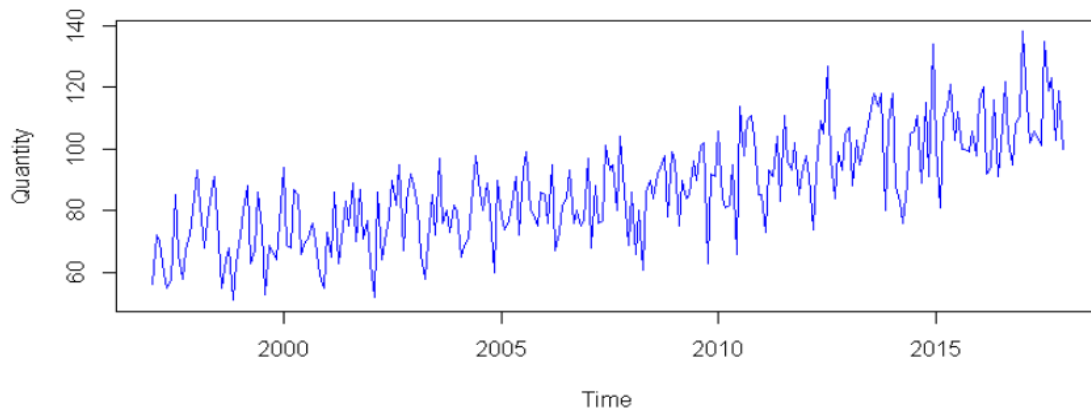


Figura 103: Evaluación - Comportamiento de la mortalidad de Arica y Parinacota

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Arica y Parinacota.

Tabla 215: Evaluación - Resumen del modelo ARIMA de Arica y Parinacota

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Arica	1.9221	11.5922	9.3648	0.4937	10.9303	0.7577

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.6446 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

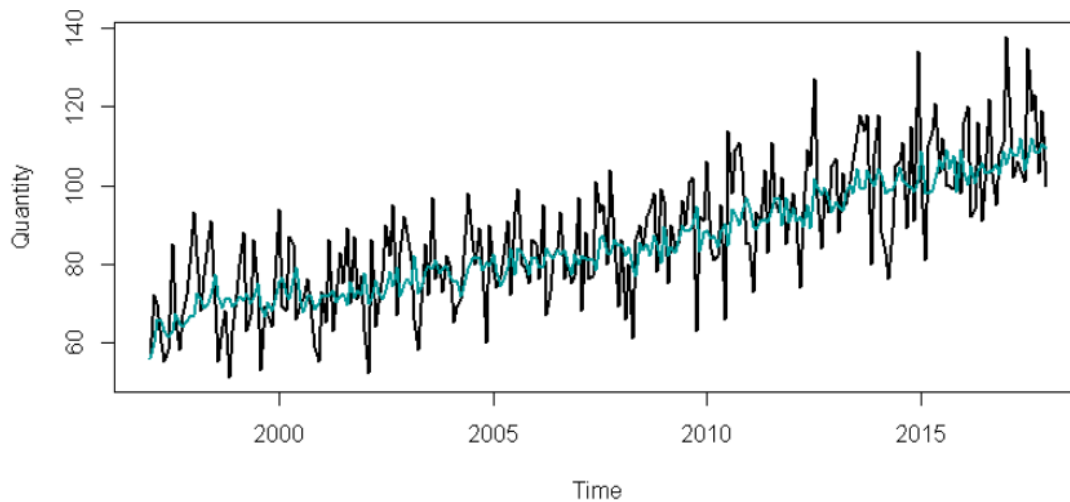


Figura 104: Evaluación - Ajuste del modelo ARIMA de Arica y Parinacota

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Arica y Parinacota.

Tabla 216: Evaluación - Predicciones estimadas para la región Arica y Parinacota

Month	Point Forecast	LO 95	HI 95
Jan 2018	114.7500	91.75415	137.7459
Feb 2018	114.5961	91.31098	137.8813
Mar 2018	107.5291	84.19828	130.8599
Apr 2018	108.2707	84.88132	131.6600
May 2018	110.8661	87.41391	134.3184
Jun 2018	107.1864	83.67132	130.7015
Jul 2018	112.8122	89.23457	136.3899
Aug 2018	114.3394	90.69932	137.9794
Sep 2018	112.3539	88.65157	136.0562
Oct 2018	108.0734	84.30906	131.8378
Nov 2018	111.9988	88.17246	135.8251
Dec 2018	109.7230	85.83492	133.6110

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

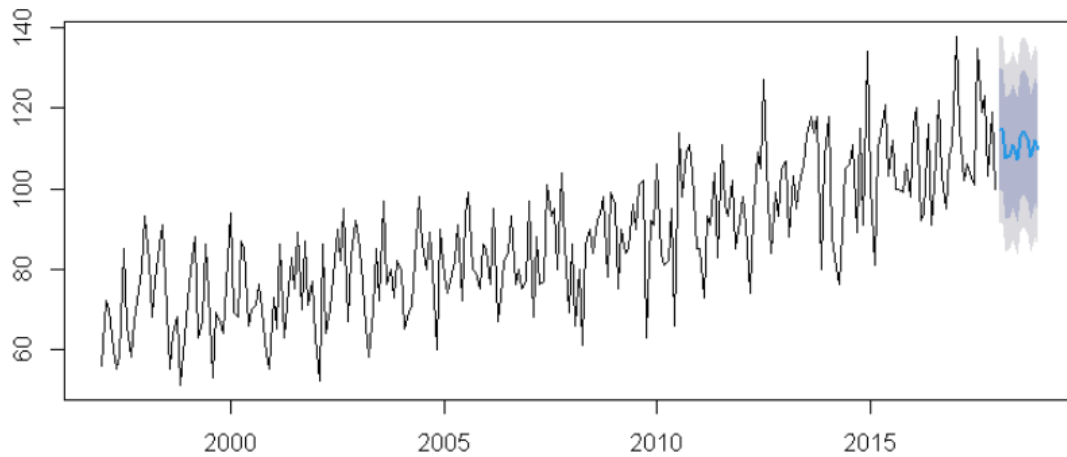


Figura 105: Evaluación - Pronóstico de la mortalidad de Arica y Parinacota

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región de Magallanes y de la Antártica Chilena

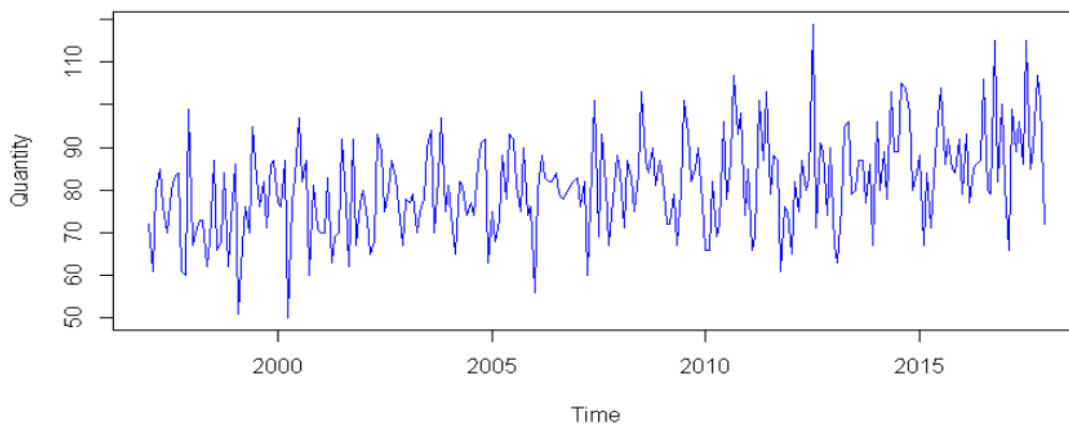


Figura 106: Evaluación - Comportamiento de la mortalidad de Magallanes

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Magallanes.

Tabla 217: Evaluación - Resumen del modelo ARIMA de Magallanes

Modelo	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Valparaíso	1.3333	10.7914	8.4811	-0.1111	10.5884	0.7343

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.8049 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

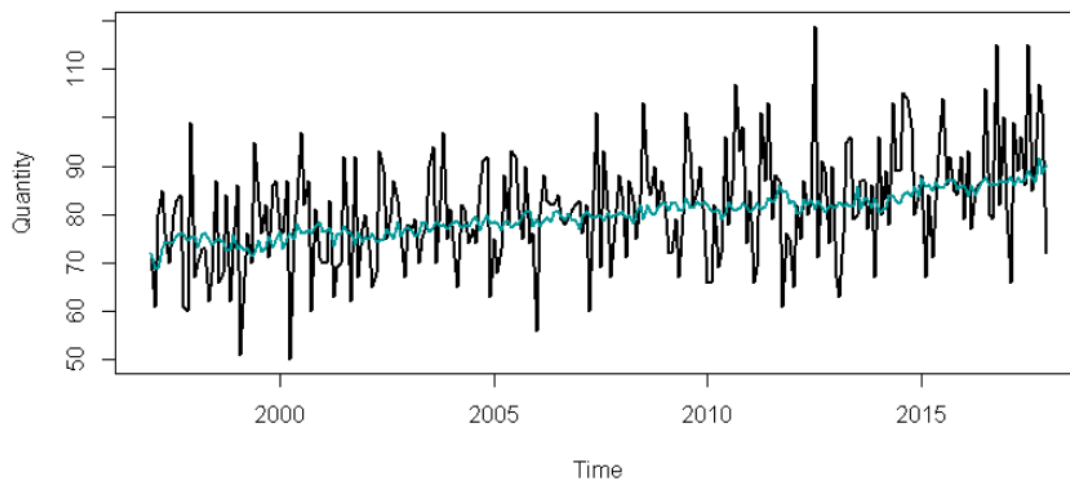


Figura 107: Evaluación - Ajuste del modelo ARIMA de Magallanes

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Magallanes.

Tabla 218: Evaluación - Predicciones estimadas para la región de Magallanes

Month	Point Forecast	LO 80	HI 80
Jan 2018	88.25969	66.89590	109.6235
Feb 2018	87.26917	65.89802	108.6403
Mar 2018	90.28361	68.91175	111.6555
Apr 2018	89.12202	67.73168	110.5124
May 2018	89.87536	68.46649	111.2842
Jun 2018	88.83657	67.41019	110.2629
Jul 2018	91.66801	70.22412	113.1119
Aug 2018	88.79343	67.33201	110.2548
Sep 2018	89.33408	67.85515	110.8130
Oct 2018	90.71049	69.21406	112.2069
Nov 2018	90.44504	68.93113	111.9590
Dec 2018	87.22481	65.69343	108.7562

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.



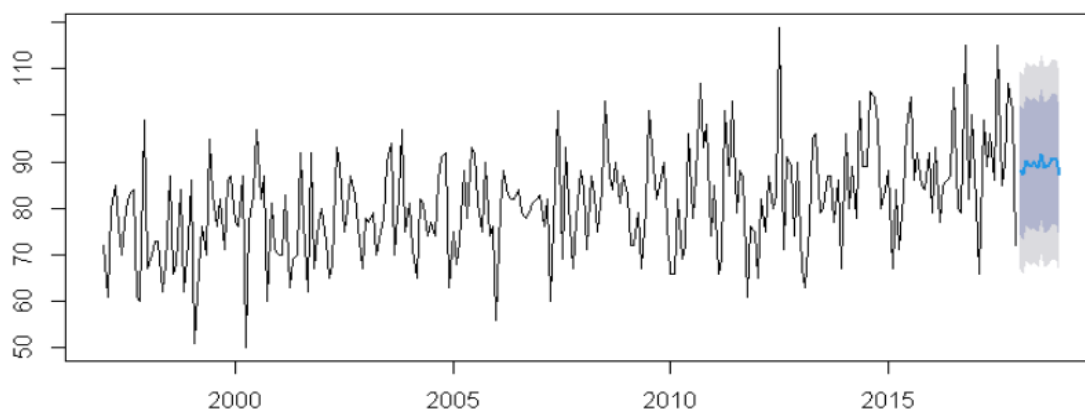


Figura 108: Evaluación - Pronóstico de la mortalidad de Magallanes

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

### Región Aysén del Gral. C. Región Ibáñez del Campo

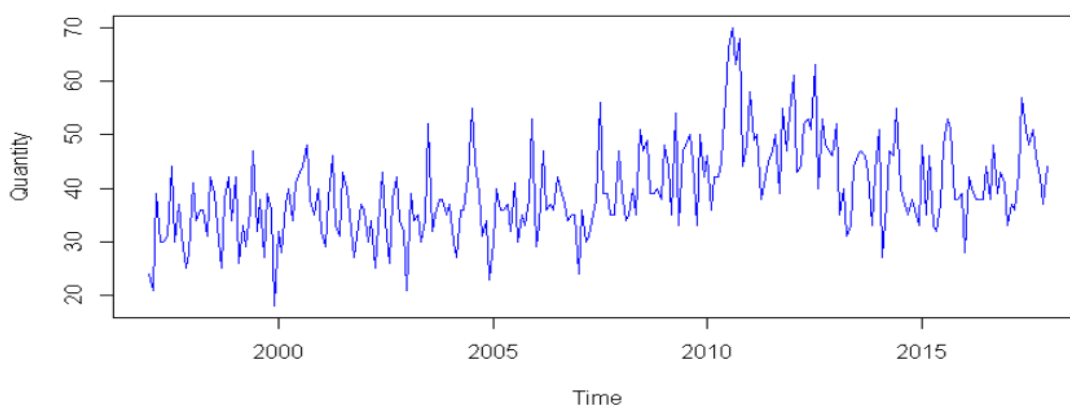


Figura 109: Evaluación - Comportamiento de la mortalidad de Aysén

Nota: Este gráfico muestra el comportamiento de las muertes en la región de Aysén.

Tabla 219: Evaluación - Resumen del modelo ARIMA de Aysén

Model	ACCURACY RESULTS					
	ME	RMSE	MAE	MPE	MAPE	MASE
Aysén	0.4768	6.9428	5.4345	-1.6048	14.1679	0.7092

En la prueba de Ljung-Boxel modelo presenta un buen ajuste, debido a que el valor es de 0.9149 lo cual es mayor a 0.05 por lo que hay ruido blanco y eso quiere decir que el modelo está correctamente ajustado.

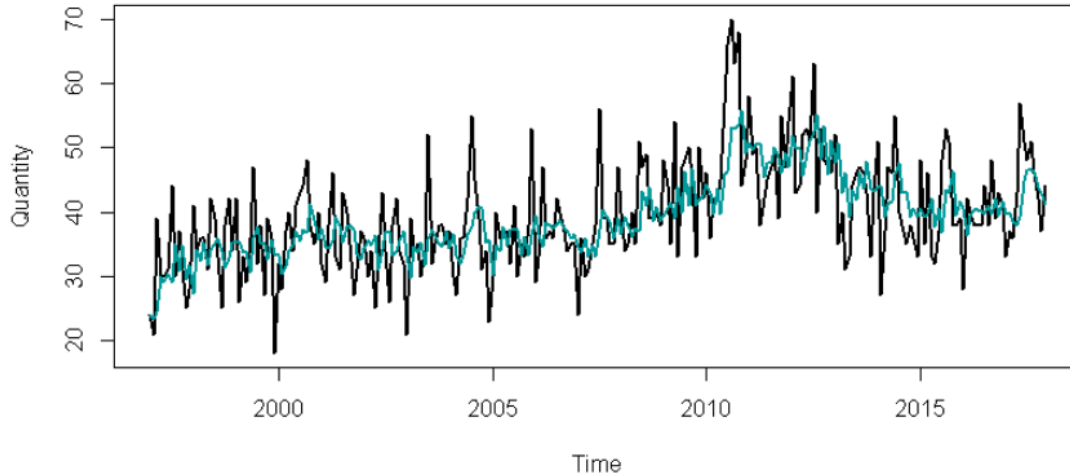


Figura 110: Evaluación - Ajuste del modelo ARIMA de Aysén

Nota: Este gráfico muestra los datos pronosticados o reales para la región de Aysén.

Tabla 220: Evaluación - Predicciones estimadas para la región de Aysén

Month	Point Forecast	LO 95	HI 95
Jan 2018	40.57655	26.74757	54.40553
Feb 2018	42.24861	27.84172	56.65549
Mar 2018	42.67111	28.06099	57.28123
Apr 2018	42.11041	27.18449	57.03634
May 2018	43.95289	28.99886	58.90691
Jun 2018	42.74975	27.62637	57.87314
Jul 2018	43.62934	28.47628	58.78239
Aug 2018	42.79001	27.52575	58.05428
Sep 2018	43.81328	28.50783	59.11873
Oct 2018	42.53906	27.14718	57.93094
Nov 2018	42.73390	27.29173	58.17607
Dec 2018	42.97025	27.45333	58.48717

Nota: La tabla indica el pronóstico para cada mes del año con un intervalo de confianza al 95%.

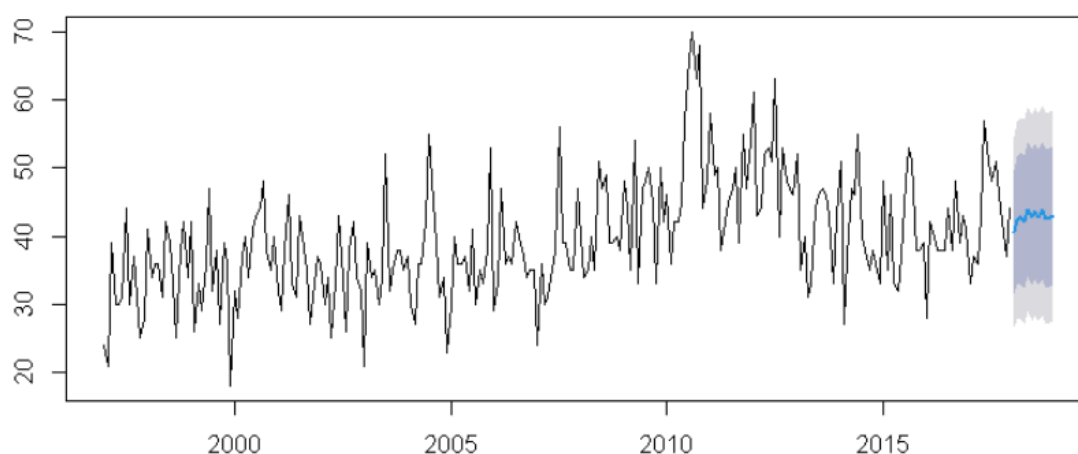


Figura 111: Evaluación - Pronóstico de la mortalidad de Aysén

Nota: Se puede ver que la línea azul es el pronóstico del modelo ARIMA y las áreas grises son los intervalos de confianza.

Tabla 221: Evaluación - Resumen de los modelos ARIMA por cada región

REGIÓN	ME	RMSE	MAE	MPE	MAPE
Metropolitana de Santiago	-5.1146	141.4304	98.4152	-0.5057	3.3578
Valparaíso	-1.8786	47.48408	35.6638	-0.5754	3.9762
Biobío	-1.1608	41.5533	30.2635	-0.5345	4.3288
Maule	8.8537	40.2790	28.8810	1.2864	5.7546
La Araucanía	-1.5589	30.2382	22.1755	-0.7386	4.5570
Libertador General Bernardo O Higgins	0.2544	26.9128	20.752	-0.4278	5.2084
Los Lagos	9.4401	29.9794	22.8530	2.0986	5.8527
Coquimbo	-1.1392	21.87694	16.5561	-1.1040	5.7282
Ñuble	-1.2505	19.2729	14.5753	-1.0451	5.7922
Antofagasta	-0.1138	23.0759	18.6217	-0.8233	8.6923
Los Ríos	2.1141	17.8165	13.8879	0.2667	7.0812
Atacama	-1.4040	12.7431	9.7610	-0.0843	9.2153
Tarapacá	1.7380	11.8725	9.2250	0.4675	9.5728
Arica y Parinacota	1.9221	11.5922	9.3648	0.4937	10.930
Magallanes y de La Antártica chilena	1.3333	10.7914	8.4811	-0.1111	10.588
Aysén del Gral. C. Ibáñez del Campo	0.4768	6.9428	5.4345	-1.6048	14.167

## 5. IMPLEMENTACIÓN

Esta es la última fase de la metodología CRISP-DM y el objetivo es explicar cómo poner en funcionamiento el proyecto que se ha construido en las fases anteriores, así como exponer los resultados obtenidos de las técnicas de minería de datos de forma didáctica e interactiva para que se pueda entender fácilmente.

### 5.1. Planificar la implementación

Los resultados de la aplicación de minería de datos se pueden visualizar didácticamente a través de un pequeño módulo con un mapa interactivo para poder consultar por cualquier región, como la curva de mortalidad, además de otros indicadores como la mortalidad por sexo, por ubicación de muerte, las enfermedades más comunes y entre otras cosas, para el desarrollo del módulo se utilizan las siguientes herramientas y marcos:

- **Marco Laravel:** Es un entorno de desarrollo de acceso libre para crear aplicaciones PHP y servicios web, Esta herramienta ayudará en el desarrollo rápido y elegante del módulo de pronóstico de muerte.
- **Wampserver:** Es un conjunto de herramientas que ofrece un servidor web Apache, una base de datos MySQL y el lenguaje de desarrollo PHP, Esta herramienta ayudará a tener un servidor para alojar el proyecto desarrollado en laravel.
- **ArcGIS Online:** Esta herramienta ayudará a alojar los mapas que se mostrarán más adelante en el módulo web y pueden ser consultados por una API.
- **ArcGIS API JavaScript:** La API de ArcGIS ayuda a tener una conexión entre la aplicación web con ArcGIS Online para intercambiar información y ver los mapas cargados en ArcGIS Online.

- **ApexChart JavaScript:** Es una biblioteca JavaScript de gráficos interactivos que servirá para visualizar los datos y tener una mejor visión del comportamiento de los mismos.
- **Plotly JavaScript:** Esta biblioteca se utilizará para proyectar las series de tiempo y el pronóstico de muertes; esta herramienta proyectará gráficos interactivos para su uso de decisiones.

## 5.2. Elaborar el informe final

En este paso, se debe presentar un informe que resuma los puntos importantes del proyecto y la experiencia adquirida durante su desarrollo. El informe final está dirigido al sector de la salud, de modo que se pueda estudiar la situación actual y se puedan tomar medidas correctivas para reducir la tasa de mortalidad. A continuación se muestra el módulo de pronóstico.

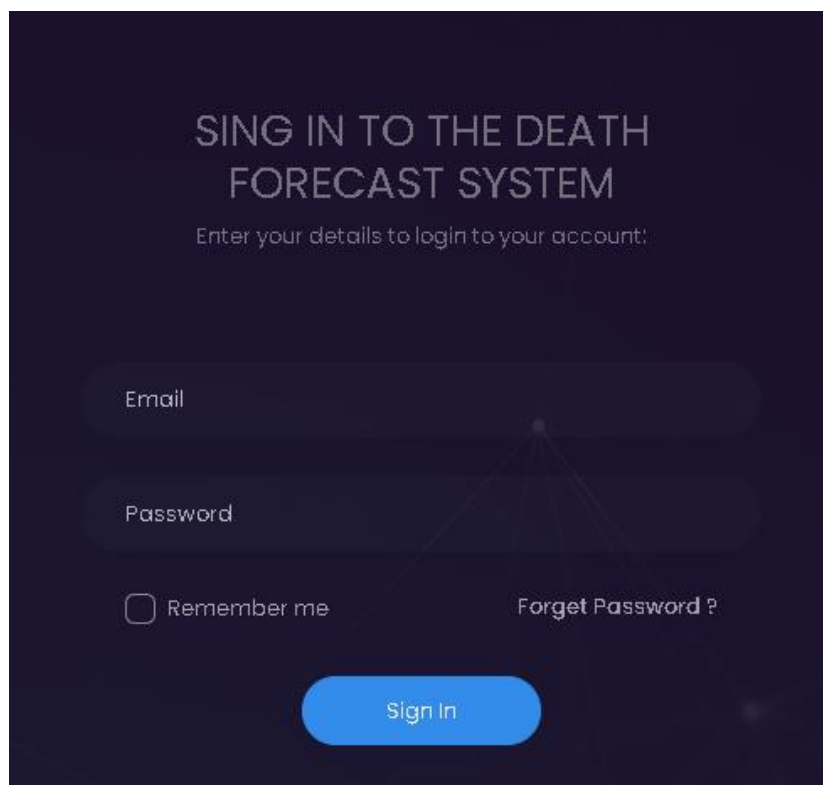


Figura 112: Implementación - Interfaz del Login

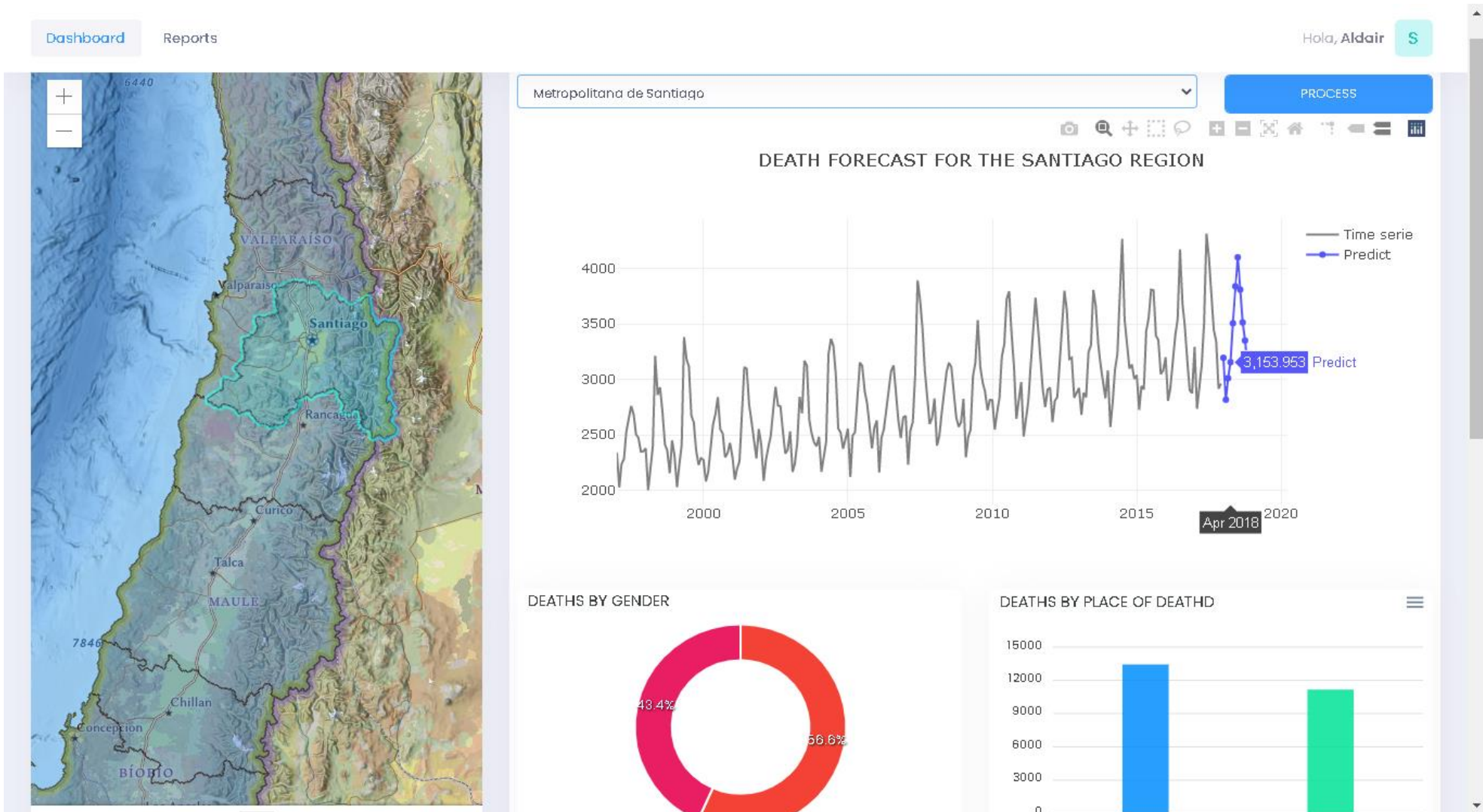


Figura 113: Implementación - Interfaz del Dashboard 1

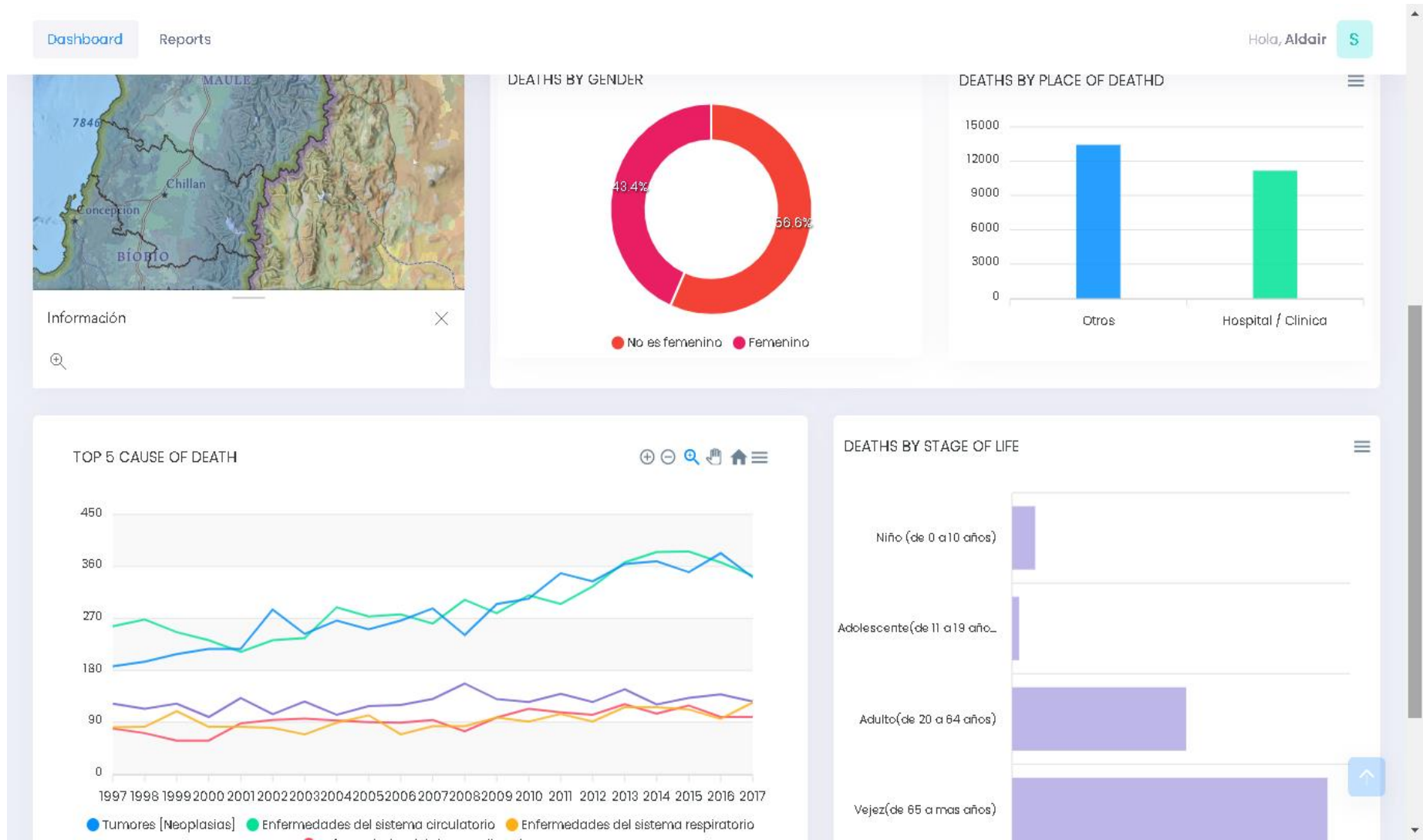


Figura 114: Implementación - Interfaz del Dashboard 2